Persistent Locators for Federal Government Publications

Summary of a study conducted for the Depository Services Program and the National Library of Canada

November 19, 2002

by
Meredith Dickison

Table of Contents

EXECUTIVE SUMMARY
Recommendations

UNIQUE IDENTIFIER SYSTEMS
A. Overview
B. Systems

1. Uniform Resource Names URNs)
2. Persistent URLs (PURLs)
3. The Handle System
4. Digital Object Identifier (DOI) / International DOI Foundation (IDF)
5. Archival Resource Key (ARK)

C. Other Tools

- LOCKSS (Lots of Copies Keep Stuff Safe)
- LinkOpenly
- OpenURL

D. Other Options

ISSUES RELATED TO PERSISTENT IDENTIFIERS / PERSISTENT ACCESS
Implications

- A. Lack of Persistence
- B. Costs and Risks
- C. Limits to Better URL Management
- D. Persistence Is Not The First Priority

- E. Timing Not Right For Managed System of Persistent Locators
- F. Leadership and Funding Needed Now

RECOMMENDATIONS

ENDNOTES

BIBLIOGRAPHY

Top of Page

Executive Summary

This study was conducted in February and March 2002 to review existing research and current practice with respect to persistent locators/digital object identifiers and to recommend steps to implement a system of persistent uniform resource locators in federal government publishing. During the interview phase the researcher undertook a comprehensive scan and review of existing research; interviewed and/or corresponded with federal government employees responsible for federal government websites; and corresponded with researchers in persistent locator initiatives. Interviewees then reviewed the draft report, made recommendations for refinement, and approved both the final version as well as its recommendations.

At present, the federal government has in place a framework to manage electronic information throughout its life cycle and has established some standards and guidelines for publishing electronic resources on the World Wide Web (the Web) that provide a basic foundation for promoting access and persistence. However, because Uniform Resource Locators (URLs), unique identifiers used to access electronic resources through the Web, are location-based, when electronic resources are moved or deleted, all links to the resource are severed and resources become inaccessible. Estimates vary, but sources indicate that the life span of a typical electronic publication lasts anywhere from 44 days to 2 years.

Through a series of interviews with individual departments, the researcher found that individually some departments follow "best practices" to manage URLs (including the use of re-directs), thereby promoting persistence, and develop holistic approaches to managing their electronic resources. However, the basic foundation appears to be in need of some additional re-enforcement. Some of the guidelines that have been developed appear to be difficult for employees to implement practically, there is a lack of consistency in implementation across departments, and some guidelines are not sufficiently rigorous or comprehensive. In addition, not all employees appear to be aware of the standards and guidelines relevant to managing electronic information.

The National Library of Canada and the Depository Services Program, Communication Canada, have unique skill sets in classifying and managing information. These skills sets should be leveraged in the arena of electronic federal publishing to ensure persistence and long-term access to valuable federal information. Persistence is not an individual institutional issue but a government-wide issue, therefore solutions need to be developed and implemented collaboratively, across government. With the federal government's commitment to implementing government online (GOL) by 2005, it is critical that the issue of persistent access is addressed now. Funding is required to ensure resources are available to put in place effective persistence solutions.

Top of Page

Recommendations:

- The National Library of Canada and Depository Services Program should make a commitment to persistence and establish a joint Working Group or Task Force to act as a vehicle for promoting persistence. The Working Group should seek to attain standing as part of the Internet Advisory Committee in order to establish a presence with key decision-makers and to develop a channel they can utilize to bring their expertise to bear on the issue.

- The Working Group or Task Force should bring together federal government managers and staff from a wide variety of backgrounds, departments, agencies and perspectives.

- The Working Group should seek to exercise influence and apply its expertise with the goal of promoting persistence in federal electronic information management through awareness raising, strengthening of federal standards and guidelines, seeking funding for collaboration, and participation in pilot projects.

The risk of not having access to information due to a lack of URL persistence is high. Public users count on government information every day and there is no way to put a dollar value on time they spend looking for information or any potential loss from not being able to find and access it. As the federal government develops larger and more complex web portals to cluster electronic information from many sources, and moves toward delivering more and more government services on-line, ensuring that links to electronic information are persistent and resources are accessible becomes ever more critical.

Top of Page

Unique Identifier Systems

A. Overview

There are essentially five formal identifier systems, all of which have been written up extensively in the literature.

1. Uniform Resource Names / Uniform Resource Identifiers (URNs / URIs)
2. Persistent Uniform Resource Locators (PURLs)
3. Handle System
4. Digital Object Identifier (DOI)
5. Archival Resource Key (ARKS)

Each of these systems is a "managed" system dependent on some type of registry structure that needs to be maintained to ensure the identifiers assigned stay unique, persistent and are updateded when ownership or other changes occur.

The limitations of URLs were recognized from the start, and hence work on developing a system of unique identifiers has been in progress for ten years. Despite this, each of the system studied has its weaknesses and strengths and all are essentially in their infancy, with the exception of PURLs.[1] In all of the systems there is independence between the "naming" system and the resolution system, which of course is in contrast

to URLs that carry their own resolution information, such as the host to contact and the protocol to use. All of the naming systems are deployed today but none widely. The reason for this is that progress in developing and deploying resolution systems that can be utilized with existing browsers has been slow.

All of the naming systems researched during this study have "umbrella" type structures. They are therefore able to encompass (additional) identification systems so nothing precludes them from being used in conjunction with other existing (publisher) numbering or identifier systems.

The key advantage or benefit of these systems (excluding PURLs) aside from providing a unique identifier is the potential for multiple resolution. The simplest resolution is the 1 to 1 resolution of, for example, an URN to an URL or to another URN. Based on a hierarchical system of resolvers, in which each resolver provides the necessary information to bring the resolution process one step further, it is also possible to turn an URN (or DOI or Handle or ARK) into multiple URLs designating the location of the resource, the address of a repository containing the resource, a bibliographic description of the resource, or an order form to obtain the resource, etc. This opens a whole new set of possibilities but also raises additional challenges when it comes to practical implementation of any of these systems, including for example:

1. Granularity - which parts of documents need to be identified and accessed separately?
2. Versioning - at what point has a document changed sufficiently to justify separate identification?
3. Copyright - do document parts with differing copyrights require separate identification?

The additional data or metadata that is involved with the identification and location of resources can be stored in the identifier string, in the locator string or in the resolution system. The metadata sets employed vary by system but generally take the approach that a limited dataset ("kernel")could be employed (in the case of ARK, as few as four (4) elements).[2]

Top of Page

B. Systems

1. Uniform Resource Names (URNs)

A naming system that includes a Namespace Identifier Code that is registered with the Internet Assigned Numbers Authority (IANA) and a Namespace Specific String, a unique and permanent code identifying the individual document. Currently there is no standard resolution system for URNs, although an Internet Engineering Task Force draft proposal exists for a resolver called the Dynamic Delegation Discovery System.[3] URNs are part of the Uniform Resource Identifier (URI) architecture, which aims to be a comprehensive and persistent resource discovery service.

A few URNs are registered and implemented (OASIS)[4] for example) but in the absence of a standard resolver they rely on techniques like catalogs that map URNs to URLs to function, on proxy servers, or on experimental resolvers.

Positives

- Naming system is flexible.

- Independent of a particular protocol and location.
- URN could refer to multiple instances of an item.
- Could be deployed in centralized or decentralized model.
- Can accommodate or adapt to any existing identifier scheme currently in use.
- Will differentiate from URL and address based schemes and implies commitment to persistence of access on the part of the assigner.
- Part of IETF/W3C work.

Negatives

- No significant deployment of URNs at this time and no standard resolution service for URN identifiers.
- Uncertain deployment prospects - W3C has been working on this issue but without much vigour.[5]
- Currently not accessible using standard browsers except through a proxy server (so identifiers have to be encapsulated in URLs using the proxy server address).
- No browser plug-in available.
- Many document formats use URNs, but some outstanding issues related to use in XML.
- Lacks support system, as there are no administration or management tools currently available.

Use of URNs in conjunction with the National Book Number (NBN) was endorsed in principle by the Conference of Directors of National Libraries. No resolution system has been defined but it would likely be through national bibliographic databases or web index or archives. In the Nordic countries this URN application is in use with resolution based on Nordic Web Index application.[6]

Top of Page

---

2. Persistent URLs (PURLs)

Functionally, a PURL is a URL, but instead of pointing directly to the location of an Internet resource, a PURL points to an intermediate resolution service. A PURL resolution service associates the PURL with the actual URL and returns that URL to the client. The client can then complete the URL transaction in the normal fashion. PURLs use standard HTTP protocols to connect to PURL resolvers and standard HTTP redirects to return information to the requesting client. Someone has to operate the PURL resolvers that provide this persistence. If the associated URL of a PURL becomes outdated, resolution of the PURL may fail. The PURL and its full history will be available as long as the PURL service itself is maintained.

PURL resolver software is available on the Web from Online Computer Library Center Inc. (OCLC). They don't keep a record of how many institutions are using their software but PURLs are fairly widely deployed. The National Library of Australia is a PURL user although their experience was not entirely satisfactory. They found there was little uptake but they recognize that it could be due to the fact that it was not widely publicized.[7]

The United States Government Printing Office (US GPO) has apparently the largest PURL installation. Since US GPO provides federal depository libraries with access to electronic documents through GPO Access, it has assumed responsibility for their

permanence although there is no explicit legal requirement to permanently maintain online electronic versions of government documents. However, the GPO believes that it has a legal basis to maintain permanent access to electronic information.[8]

Positives

- A simple solution using standard HTTP redirects and URLs so no major training required.
- Includes a collection of utilities that support the management of URLs by allowing users to update their own resource, a critical aspect of assuring the integrity and maintenance of links.
- Software readily available from OCLC.
- Many organizations could deploy a PURL server so a centralized or decentralized model could be deployed.
- Operating a PURL would signal a commitment to the persistence of the resource.

Negatives

- Not a standard; only a technique.
- When the current publisher ceases to provide access services, there is no plan for forwarding.
- No object descriptions or commitment statements required.
- Not location independent.
- Protocol dependent.
- Can only support a one-to-one resolution.
- Standard browsers return the URL in the location field. (This would in many cases become the citation unless the PURL is clearly indicated on or in the object itself or browsers are modified to display the persistent URL that was requested rather than the actual URL representing the current location of the resource.)
- Almost always seen as an interim solution not a final answer to persistence.
- Some specific software problems.[9]

Top of Page

---

3. The Handle System

A distributed persistent naming system developed by the Corporation for National Research Initiatives (CNRI). The system manages "handles" which are unique names for digital objects and is supported by a resolution system consisting of a distributed system of global, local and caching servers. A Global Registry Handle Registry run by CNRI registers top level naming authorities.

The Handle System is implemented by the Library of Congress Digital Library.[11] Another implementation is D-Space, a joint MIT Library and Hewlett-Packard project housing 10,000 articles produced annually by Massachusetts Institute of Technology faculty and researchers.[12] Also implemented by International DOI Foundation - details on the next page.

CNRI is only aware of one Handles System implementation in Canada.[13]

Positives

- Proven - already implemented in major applications; interoperable with DOI if that system becomes developed widely.
- Flexible naming system, URN compliant.
- Built-in resolution mechanism.
- Able to be deployed in a centralized or decentralized model.
- Independence of a particular protocol.
- Allows for multiple resolution.
- Software and management tools are available from CNRI.
- Developed specifically for digital library applications.
- Can accommodate existing schemes easily.
- May be used in tandem with LDAP or WHOIS++ for reverse lookup service.

Negatives

- Not simple to establish.[14]
- Not yet useable through standard browsers without plug-in software for the client.
- Not registered as a URI/URN scheme (procedures for doing so are unclear and relative advantages of registration as either a URI or URN not clear). However, unlikely to be implemented by standard browser developers if not formally recognized by W3C.
- CNRI as a research initiative raises the question of support for system if funding is withdrawn and the body ceases to exist although given investment of LC and the IDF unlikely the system will suddenly die).[15]
- Some security limitations.[16]
- Complex, proprietary resolution infrastructure required for what is really a straightforward table lookup.
- No object descriptions or commitment statements required.

Top of Page

---

4. Digital Object Identifier (DOI) / International DOI Foundation (IDF)

This is a system managed by the International Digital Object Identifier Foundation with three elements:

1. A persistent identifier of intellectual property along with syntax and assignment rules.
2. A resolution system using the Handle System technology.
3. A metadata structure to identify the items represented by the DOI (based on the INDECS metadata set).

DOIs are assigned by Registration authorities (currently there are four, three in the US and one in Korea) and registered in the Global Registry.[17]

The business model is in a state of flux,[18] but IDF is intended to operate on a cost recovery basis. Governance issues are also not settled at this time.

Although 200 organizations[19] are allocating DOIs and over four million DOIs have been issued as of January 2002, views in the literature are mixed. On the plus side, it is being implemented by the U.S. Department of Defense.[20] Others say the system is not proven and "may fizzle", although there is no comparable technology on the horizon.[21] Although developed out of the specific needs of the publishing industry, not all

publishers have embraced it and it is evolving to meet broader user needs.[22]

In Canada, NRC Research Press is assigning DOIs to journal articles.[23] They pay $750 US for a membership plus 60 cents a DOI to deposit and 10 cents to retrieve, but retrieval costs can be eliminated through local caching. They publish the DOIs in the printed copies of their journals. Their Web Site also provides instructions for linking to articles with DOIs (i.e. by using CrossRef) as well as URL links (aliases) to ensure stable, long-term linking to individual journal homepages. They are satisfied with their experience with DOIs to date and have found it easy to maintain (essentially a file with the assigned DOI numbers and metadata are sent to CrossRef monthly). They see DOIs as one vehicle for achieving persistence but they are also using DOIs to get broader exposure in the marketplace.

Positives

- Fully managed system supporting centralized or decentralized models.
- Flexible naming system can accommodate any existing schemes easily.
- Has a built in resolution infrastructure.
- Independent of a particular protocol.
- Allows multiple resolution.[24]
- Supported by freely available software (many of the same advantages of the Handles system).
- System may at a minimum be widely deployed by large commercial publishers.
- Institution could become a Registration Authority that offers potential to build value-added services.
- System can be further development in an open standards environment.
- IDF has begun to integrate into related standards activities for JPEG (involvement with Rights data dictionary project).
- IDF is beginning to get support from technology companies like Adobe Acrobat and Microsoft.[25]

Negatives

- High risk in the sense that system is not proven - it remains to be seen whether it will be widely deployed in the long term.
- Cost to move to a totally managed system; also many uncertainties about costs at this time.
- Many uncertainties about the extent to which the interests of large publishers will prevail; longer term, costs may be prohibitive for smaller publishers and website owners; need to ask: will large publishers restrict access in future?
- Many uncertainties about governance issues associated with the establishment of a DOI registration agency (rules on registration are now relatively loose) so progress of deployment is unclear.
- Not yet useable through standard browsers without plug-in software for the client (although proxy servers could be used).
- Would be expensive to become a Registration Authority (start up fees) and of dubious value without an e-commerce game plan.
- IDF has shown only peripheral interest in libraries and archives and their needs.
- Mandates use of the Handle system for DOI resolution.
- Fee-based registration.
- No object descriptions or commitment statements required except as part of the centralized, fee-based CrossRef system.

Top of Page

---

5. Archival Resource Key (ARK)

"A founding principle of the ARK is that persistence is purely a matter of service." - John A. Kunze

Designed for custodians of archived digital objects, the ARK comprises identifier syntax and three services which give users a link from an object to a promise of stewardship for it, to a description of it and a link to the object itself. Electronic Resource Citation (ERC), one method under development for delivering ARK services, is a metadata format using four elements derived from the Dublin Core. The ERC provides an ultra simple and compact way to create and process object descriptions.[26] Compared to XML, ERC format is weak when it comes to nesting fully transparent records more than a few levels deep.

Implementation experience is limited to prototype ARK service at the U.S. National Library of Medicine, [27] the World Intellectual Property Office, and the California Digital Library,[28] A developer of ARK **reminds us** "The bar for providing ARK service is high."[29]

Positives

- Designed specifically for digital archival materials.
- Location and protocol independent.
- Can be used with standard browsers today.
- Can support multiple resolution.
- Squarely addresses the service issues behind persistence.
- Clearly separates the role of original name assignment from the role of ongoing support for access and description.
- Resolver discovery, when it becomes necessary, can be implemented with a small Perl script that uses the existing DNS system to locate the current service provider.
- Associated services are based on metadata (could recycle existing metadata).

Negatives

- Too new to judge the likelihood of acceptance - it may be too stringent for many.
- Like the URN and DOI, eventually requires a resolver discovery infrastructure - no guarantee that this will be fully developed or accepted.
- Too new to realistically assess costs of implementing such a system but undoubtedly would require considerable investment.

Top of Page

C. Other Tools

LOCKSS (Lots Of Copies Keep Stuff Safe)

Another identifier tool discussed in the literature is LOCKSS, defined as a Permanent Web Publishing and Access System.[30]

LOCKSS is a tool for libraries to ensure their community's continued access to web-published scientific journals. It allows libraries to run web caches for specific journals. The caches collect content as it is published and are never flushed. They cooperate in a peer-to-peer network to detect and repair damaged or missing pages. The caches run on generic PC hardware using open-source software and require almost no skilled administration, making the cost of preserving a journal manageable. The system was tested in 2000 at more than 40 libraries worldwide with the support of more than 30 publishers, and it proved to be technically feasible. Further testing is currently in the works (scheduled for completion August 2002),[31] but deployment prospects are unknown at this time.

Top of Page

LinkOpenly

Another open source software called LinkOpenly[32] is being developed by Openly Informatics Inc. in concert with the CrossRef DOI to provide persistent and comprehensive linking to e-journals from web pages, libraries and databases. Openly Informatics Inc. has already developed OpenlyJake[33] which provides another linking mechanism by managing online resource metadata with a database union list, title authority control, and linking tools, as well as making it easy to add a local holdings layer.[34]

Top of Page

OpenURL

The United Kingdom library community is also doing work in this area. Based on their work [35] the OpenURL provides a mechanism for encoding a citation for an information resource, typically a bibliographic resource, as a URL. The OpenURL is, in effect, an actionable URL that transports metadata or keys to access metadata for the object for which the OpenURL is provided. The target of the OpenURL is an OpenURL resolver that offers localized services in an open linking environment. The OpenURL resolver is typically referred to as the user's Institutional Service Component (ISC). The remainder of the OpenURL transports the citation. The citation is provided by either using a global identifier for the resource (for example a Digital Object Identifier) or by encoding metadata about the resource, for example, title, author, journal title, etc., or by some combination of both approaches. It is also possible to encode a local identifier for the resource within the OpenURL. In combination with information about where the OpenURL was created, this allows software that receives the OpenURL to request further metadata about the information resource.

Committee AX, the National Information Standards Organization (NISO) standards committee recently adopted an OpenURL draft as Version 0.1 of the standard. The Committee acknowledges that OpenURL technology is in its infancy, but is very optimistic that its application to bibliographic citations only scratches the surface of what is feasible. In the near future, the Committee expects third parties may provide OpenURL links for any information resource. They also expect that in the near future OpenURL links will be encoded in XML. [36]

Top of Page

D. Other Options

Aside from the options of adopting an existing identifier systems or developing and managing one's own (the approach being taken by Australia's National Library), a third option for achieving persistence may potentially exist using RDF/XML as the common resource description framework or foundation for coding and exchanging data.[37] Building on the Dublin Core metadata elements already required for government Web resources, there might be an opportunity to include an identifier that could be used to uniquely identify individual electronic publications. Further investigation is required to see how this could be done, in what time frame and in the context of already existing identifiers (GILS for example).[38]

Top of Page

Issues Related to Persistent Identifiers / Persistent Access

Another area of debate and discussion in the literature is the relationship between persistence of identification and persistence of access. The persistent identifiers previously discussed provide a means for persistent identification and persistent access as long as an item is online and the system is maintained, but do institutions really want to "guarantee" that all items with persistent identifiers will persist forever online?

A critical part of persistence is establishing a clear definition that allows an institution to clearly communicate the level of commitment it is making to maintaining the availability (access to) of a resource online. (A few institutions have a Persistence Policy posted online - see for example W3C's Site.)

The U.S. National Library of Medicine appears to have done the most practical work in this area defining three (3) core categories of permanence for electronic resources, although they have not yet actually been assigned to any resources yet [39] and in subsequent re-iteration, have been able to develop a condensed version.[40] The U.S. National Library of Medicine categories are:

1. **Identifier validity** defined as the extent to which a user can be assured that a given identifier will not change. In this category, there are two (2) options:

   (a) **undefined or transient** - meaning no rating is undertaken or the identifier could be changed or reassigned or

   (b) **guaranteed** - meaning the identifier will not be changed or reassigned.

2. **Resource availability** defined as the extent to which a given resource will remain accessible. In this category, there are two (2)options:

   (a) **no guarantee** - the resource may become unavailable at any time or

   (b) **permanently available** - meaning accessibility is guaranteed implying an archival commitment.

3. **Content invariance** or the extent to which the content of a given resource and the links it contains will remain unchanged. In this category, four options are available:

   (a) **unrated** nbs;- meaning no rating has been undertaken and no guarantee has been made

(b) **dynamic** - meaning the content may be replaced, corrected and revised and links could change

(c)**static** - meaning the content is subject only to correction and minor additions and internal links will be updated, and

(d) the content is static and will not change.

In many cases, it is possible to assign ratings to resource categories (for example, announcements, annual reports, etc) so there would be no need to assign ratings to individual resources.

The National Library of Australia has also done work in this area. In their view, information that needs to be managed for persistent access will generally fall into two types:

(a) **Static resources** - items that have a definitive version and will not be altered.

(b) **Dynamic resources** - materials that undergo change, either incremental change by addition of content or change by modification, deletion and insertion of content.

They can be either **archival** (material that you wish to preserve and provide persistent access to) or **ephemeral** (material not likely cited in other resources but frequently book marked by users looking for the latest version so this category needs persistence of citation only for the current version.)

In their model, the decision about what level of persistence of access to provide will be a matter of judgement, and this should be determined at the point of creation of the document.[41]

Top of Page

Implications

A. Lack of Persistence

- URLs are not persistent - URLs get changed for a wide variety of reasons, such as when departments change, when employees decide to re-organize file structures, when pages are removed from the Web, etc. - and links can become broken. Lack of persistence in URLs can occur in any department and this makes it a government-wide issue, not just a library/depository issue. Statistics documenting the problem could be collected since departments check links and maintain statistics on 404 error messages. There may be some merit in verifying the number of broken links by department, but the statistics would not be helpful in defining what is an acceptable level of broken URLs. Would zero tolerance be acceptable or would 2-3% be acceptable? The answer may be found in understanding the rationale for achieving persistence in the first place: the goal is not to maintain a particular URL for its own sake, but rather to ensure that a particular electronic document located at that URL can in fact be found at that URL and accessed, as long as the electronic file is on the Internet. The issue of whether electronic records should be permanent forever (on the Web or somewhere else in their original format or some other format) is germane to information management, but it is not part of this report.

Top of Page

B. Costs and Risks

- There are costs associated with checking links to promote persistence and these costs are being incurred regularly in all government departments and agencies. There are also costs incurred when links become broken. These costs are more difficult to quantify because there is really no way to keep track of how much time individual government employees spend trying to locate documents. However, even assuming the federal government as a whole and some departments in particular have a superior track record at maintaining URLs than the Internet community in general, the costs are hardly inconsequential. True, one person spending a few minutes trying to locate a file may be inconsequential. However, spread across the entire federal government bureaucracy, those few minutes can quickly calculate into hours especially as more and more federal employees are "knowledge workers" who spend all of their time working with information. It is also important to keep in mind that more and more government information is being made available on the Web and that federal government portals are growing in size and complexity.

- There is no way to calculate the costs to (public) users who require information and are unable to find it due to broken links. The public needs access to government information to make their lives better and to know of potentially harmful situations and events - bad weather, air pollution, faulty tires, to name just a few. Businessmen, scholars and researchers need timely and ready access to information to make key decisions. There is no way to put a dollar value on their time looking for the information or any (potential) loss from not being able to access the information. Even in cases when the result is only minor inconvenience and annoyance, the net effect over time is erosion of trust and confidence in government.

- Direct costs and loss of productivity are not the only problems caused by lack of persistence in URLs. There are bigger risks associated with lack of persistence, namely - the ability of the government to make effective decisions, when the government cannot efficiently access and use its own information and the ability to deliver a meaningful level of accountability. Managing URLs is only one small part of information management, but it is the individual parts taken together that determine the ability of the government to disclose and provide access to the information they have created which is the very basis of the Canadian way of life. Individual incidence of lack of access whether through ignorance, error or intent equally has a negative effect and cumulatively is corrosive to the entire system.

- Federal employees cannot fulfill their role as stewards of information without leadership, access to tools they understand and can use and proper training.

Top of Page

C. Limits to Better URL Management

- Although undoubtedly URLs can be better managed (at least in some parts of the federal government) to promote persistence, they will never provide a unique identifier independent of location so they will always pose challenges to ensuring access.

- Despite the development of global unique identifier systems, the Domain Name System (URLs) gives every appearance of prevailing as the main vehicle for accessing and locating web resources in the short to medium term.

- From this perspective, it would be useful if federal government URLs could be made "neutral" so that there would be less reason to change them, but this is practically unworkable for a number of reasons. Developing and employing consistent names and naming conventions at the sub-domain level would be beneficial by adding consistency. At the same time, there is recognition of the need to balance and reflect all stakeholder needs: user needs (for simplicity), developers needs (workable for coding and ease of maintenance) and departmental/governmental needs (marketability and branding).

Top of Page

D. Persistence is not the First Priority

- Federal government employees individually recognize the problems caused by the lack of persistent URLs but the concept of persistence does not appear to be a top of mind issue or priority for departments/government as a whole. Finding and implementing solution to the problem will first require further education about the problem and acceptance that it is a priority. In the existing decentralized distributed framework, it will take time to move this issue to the top of the Agenda unless institutions such as the National Library of Canada and Depository Services Program who have an important role in maintaining and distributing electronic publications are willing to assume a leadership role. The increasing importance of good information management (including managing URLs) calls for much greater collaboration and co-ordination among central agencies with some direct responsibility for information management, such as Treasury Board, National Archives, Clerk of the Privy Council, Auditor General, National Library and Department of Justice to name just a few.

Top of Page

E. Timing Not Right For Managed System of Persistent Locators

- A shift away from the current practice of relying on URLs to a totally managed system of persistent locators would involve significant start-up costs in time and resources and require additional on-going investment for annual maintenance. Adoption of any of the existing managed persistent identifier systems at this time is medium/high risk due to their immaturity. At this time, it is unclear if any of the existing identifier systems will ever be widely deployed. It is unclear if standard browsers will be re-configured to resolve any of these systems, without the use of plug-ins.

- Timing is another factor. It is premature to attempt to move to a fully managed persistent locator system without laying the appropriate groundwork. At this time, there is unlikely to be widespread support for such an approach within the federal government. Without a consensus,achieving results in the short to medium term by endeavouring to move in that direction is highly unlikely and a waste of already scarce resources. However, existing persistent locator systems do merit close monitoring to determine if any will achieve widespread use on the Web and to stay abreast of developments in resolutions mechanisms, especially with respect to capabilities in standard browsers to resolve URNs. Making the federal government community as a whole more aware of these developments will bring attention to the deficiencies in URLs and help promote persistence.

- The government has a general framework on which to build persistence and ensure access and persistent locators are only one aspect of access. Any decision regarding the implementation of persistent URLs should be done in the context of the lifecycle management of web resources, web content management standards and delivering GOL. As such, it also necessarily needs to be in collaboration among partners. The National Library of Canada and Depository Services Program should seek out opportunities to work collaboratively with others on this matter and participate in existing

projects that advance government wide solutions and/or develop a project that has government wide application.

- Until such time as URLs become widely displaced by URNs or DOIs or some other unique global identifier or technology, promoting persistence and access can be pursued on a number of fronts, including education on the importance of persistence and its role in delivering GOL; managing URLs better; promoting the archiving of electronic files as one critical step in the lifecycle management of electronic resources and actively collaborating with other individuals and departments charged with information management responsibilities.

Top of Page

F. Leadership and Funding Needed Now

- Although broken links and inaccessible electronic files are not the kind of attractive issue managers like to embrace, the NLC and DSP, with their depth of experience in managing information, are well positioned to show leadership on the issue. To that end, under the GOL initiative, the National Library of Canada and Depository Services Program should seek funding to obtain the resources necessary to undertake initiatives that advance the agenda and participate and/or initiate pilot project(s) to promote persistence for federal publications in electronic format. There may be an opportunity to participate in projects currently being developed by Treasury Board. The NLC and DSP should also proactively seek out other appropriate collaborative opportunities before launching their own initiative(s).

Top of Page

Recommendations

- **Recommendation One**
  The National Library of Canada and Depository Services Program make a commitment to persistence and establish a joint Working Group or Task Force to act as a vehicle for promoting persistence. The Working Group should seek to attain standing as part of the Internet Advisory Committee in order to establish a presence with key decision-makers and develop a channel they can utilize to bring their expertise to bear on the issue.

- **Recommendation Two**
  The Working Group or Task Force should bring together federal government managers and staff from a wide variety of backgrounds, departments and agencies and perspectives.

- **Recommendation Three**
  The Working Group should seek to exercise influence and apply its expertise with the goal of promoting persistence in federal electronic information management through:

  1. Awareness and promotion

     - Promote awareness and understanding of the problem and position it as a technological and social issue.
     - Promote the NLC and DSP as partners with federal departments and other

agencies in supporting persistence and access to government electronic resources.

- Promote better URL management and the sharing of information on "best practices".
- Promote the adoption of a statement on persistence.

2. Strengthen Federal Standards and Guidelines

- Enhance the usability of existing relevant federal standards and guidelines so they are implemented consistently across the government and employees have the knowledge and tools to be good stewards of electronic information.
- Strengthen relevant federal standards and guidelines by filling gaps and making revisions and updates as appropriate. For instance, develop a communications standard and modify the guidelines with respect to vanity URLs; consider establishing guidelines for sub-domain and sub-directory file names.

3. Seek Funding for Collaboration and Participation in Pilot Projects

- Investigate opportunities to participate in pilot project(s) that promote persistence, have government-wide applicability and enhance the delivery of federal government services on-line. If an appropriate project is available [47], seek funding under the GOL initiative to participate; if not, develop a business case to obtain funding for a project jointly initiated by NLC and DSP. It is recommended that the NLC and DSP consider a pilot project testing the concept of a national repository for select electronic information based on a physically decentralized model and a unique identifier system.

Top of Page

ENDNOTES

1 PURLs is unique in that most of the literature discusses it as an "interim" solution to persistence. One exception stated that: "A commonly held belief is that the PURL Resolution Service is a predecessor of the URN-resolution system and that once all the URI-standards are in place, PURLs will become obsolete. However a PURL system is a URI-resolution system that provides for the simple one-to-one resolution of a PURL to a URL. This type of simple resolution for URL-management purposes will remain important for certain types of web-resource collections, now and in the future. It is very likely that complex URN-resolution systems will emerge in the future to control access to resources of specific (commercial) resource collections. It is even likely that these complex systems will be more manageable if location-resolution goes through a PURL-like system, maintained separately. The functional requirements for URL-management and those for other access-parameters may justify the implementation of distinct systems." See document entitled "Identification, Location and Versioning of Web-Resources".

2 For details see article in bibliography entitled "A Metadata Kernel for Electronic Permanence".

3 For details refer to draft standard "Dynamic Delegation Discovery System (DDDS)" by M. Mealing in the bibliography.

4 Organization for the Advancement of Structured Information Standards at www.oasis-open.org.

5 See comments in Australian Report entitled "Persistent Identification Systems", Part 1, section 1.3.4.

6 For further information see the articles by J. Hakala and registration of namespace at Helsinki University website listed in the bibliography.

7 See their Report "Persistent Identification Systems".

8 GAO Document 01-428, "Information Management: Electronic Dissemination of Government Publications, Volume 1", March 2001.

9 See Web site www.purl.oclc.org/problems.html.

10 See some of their documentation in the bibliography in the Handle System Section. For information on naming see document on their web site entitled "LC Naming Conventions for Digital Resources".

11 Refer to article in the bibliography entitled "A Test Bed for Digital Object Registration and Management: The Copyright Office Electronic Registration, Recordation and Deposit System: 2000 Project Summary".

12 See article in bibliography entitled "Some Applications of the Handle System".

13 The location is a private security business according to an e-mail from Jane Euler dated January 30, 2002.

14 Australia for example did a pilot test of The Handles System (see Appendix 2 of their Report entitled "Persistent Identification Systems"). They found that while the system has the basic technical functionality for a digital library and is the best option currently available, they could not recommend it because of the investment required for the establishment of the system and procedures and the system modifications required. The system also requires UNIX and Java skills of a high order.

15 IDF negotiated guaranteed levels of support and service level standards with CNRI, who have moved support for the system to a commercial agency and now guarantee a 24 x 7 operation.

16 See article in bibliography entitled "Overview of Handles System".

17 CrossRef serving the journal publishing industry; Content Directions and Enpia Systems (Korea) serving the e-book sector and Learning Objects Network which intends to register DOIs for use in the management of learning objects.

18 Business model is in a state of flux, but in original model the cost is $1,000US for a prefix which allows for any number of DOIs to be issued. In future, it is expected that the price per prefix will vary depending on the Registration Authority and the range of services each offers.

19 See example of DOIs as used by Academic Press' IDEAL system at www.apnet.com

20 See article entitled "The DOD Does DOI: Massive Project May Put Wheels Under DOI Initiative".

21 See article entitled "DOI: A Tale of Whoa" in bibliography.

22 See its applicability to reference linking in article in bibliography entitled "Reference Linking With DOIs: A Case Study".

23 At this point in time, they do "linking in" but not "linking out", a feature designed to provide users with links to other relevant web-accessible resources.

24 For example, CrossRef working with a number of other organization has

demonstrated the ability to do "Localized Linking" which allows researchers to be directed to a version of an article that provides full text access without payment (say through a local library subscription) rather than to a version where payment would be required.

25 Microsoft has decided to support DOI in its eBook reader. Some demonstration scenarios at www.seyboldreports.com/ebooks/features/011010-doi.html.

26 According to an e-mail from J. Kunze, ARK's developer. (E-mail dated October 2, 2002)

27 According to an e-mail from J. Kunze, ARK's developer, at this stage implementation is limited to the development of about six (6) pages of PHP code that uses the Entrez API to the Pubmed database which illustrates the protocol and metadata format described in the ARK specification. "The hard work of assigning permanence ratings has not been done". (E-mail dated January 24, 2002)

28 According to an e-mail from J. Kunze. (E-mail dated October 2, 2002).

29 See John Kunze article in bibliography entitled "A Metadata Kernel for Electronic Permanence".

30 All of the available information on this tool is listed in the bibliography.

31 According to an e-mail from Victoria Reich, Project Manager dated January 27, 2002.

32 For details see their home page at http://www.openly.com

33 Also being used by NRC Research Press.

34 They are based on OpenURL technology which is a resolution service that considers local user information when directing users to a resource. For more information see articles in bibliography. DOI considers OpenURL technology to be complementary. According to NISO, at this point in time, it is not clear how the proposed OpenURL standard should relate to metadata standards and identifier standards.

35 See article in bibliography by Andy Powell entitled "Encoding OpenURIs in Dublin Core Metadata".

36 NISO Press Release dated February 11, 2002. For more information refer to http://library.caltech.edu/openurl/
default.htm

37 See W3C standards for RDF/XML. See also their standards for XPath, a software for addressing XML documents.

38 There are few specifics in the literature on how this approach could actually be applied. Some German scholarly scientific publishers appear to have done some work in this area, but details are unavailable at this time since all source documents located were the German language only (see CARMEN project at www.bis.uni-oldenburg.de/carmen_ap4/). The UK library community has also looked at RDF in the context of their community. See UKOLN RDF Seminar notes on their Web Site www.ukoln.ac.uk.

39 According to an e-mail from the Library's Chair of the Working Group on Permanence of Electronic Information, M. Byrnes dated January 25, 2002.

40 The condensed version is available in the article in the bibliography entitled "Defining NLMs Commitment to the Permanence of Electronic Information", p. 3.

41 See their document entitled "Safeguarding Australia's Web Resources: Guidelines for Creators and Publishers" listed in the bibliography.

42 Preliminary information suggests there may be an opportunity to participate in a metadata-based Treasury Board initiated project, but project details are not available at this time. The Working Group should proactively pursue this opportunity. If it is not appropriate, the Working Group should seek to identify another opportunity, before proceeding to initiate one.

Top of Page

Created: 1996-05-22
Updated: 2004-01-12

Top of page **Important Notices**

Persistent Locators for Federal Government Publications

Bibliography

Please note: this partially annotated bibliography was compiled in February and March 2002. Some links may no longer be active.

- General
- Persistent Identifiers - Specific Systems
  - The Handle System
  - DOI System
  - ARKs
  - PURLs
- URLs, URNs and URIs
- URNs and NBNs
- Other Namespaces
- Publisher Identification Systems
- Persistence in Journal Citations
- Australian Resources on Persistence
- Organizations
- OpenURL
- Miscellaneous

Report - Table of Contents

General

Arms, William. "Key Concepts in the Architecture of the Digital Library". D-Lib Magazine. July 1995. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/July95/07arms.html

Caplan, Priscilla and Flecker, D. "Choosing the Appropriate Copy". Digital Library Federation Architecture Committee. September 1999. Accessed on March 5, 2002.
URL: http://www.niso.org/DLFarch

Causton, Laurie. "Identifying and Describing Web Resources". European Commission DGXIII/E-4. Accessed on March 5, 2002.
URL: http://www.elpub.org/html/webres.html

CrossRef : The Central Source for Reference Linking. Accessed on March 5, 2002.
URL: www.crossref.org/

Doyle, Mark. "Citing and Linking in Electronic Scholarly Publishing: a Pragmatic Approach". American Physical Society. [1999?]. Accessed on March 5, 2002.
URL: http://ridge.aps.org/DOYLE/sweden/talk/sweden.html

Engelbart, Douglas C. "Knowledge-Domain Interoperability and an Open Hyperdocument System". June 1990. (AUGMENT,132082,). Accessed on March 5, 2002.
URL: http://www.bootstrap.org/augment-132082.htm

Flecker, Dale. "Citation - E-Journal Links from the (Academic) Library Viewpoint: A presentation to the NISO/NFAIS/SSP Linking Workshop". February 11, 1999, Washington, DC. Accessed on March 5, 2002.
URL: www.niso.org/presentations/flecker.html

Green, Brian and Bide, Mark. "Unique Identifiers: A Brief Introduction". [1997]. Accessed on March 5, 2002.
URL: www.bic.org.uk/uniquid.ht

Guenther, Rebecca. Naming Conventions for Digital Resources. Library of Congress Network Development and MARC Standards Office. Accessed on March 5, 2002.
URL: http://lcweb.loc.gov/marc/naming.html

Information Object Numbering Systems (IONS): JTAP report 42. Division of Learning Development, De Monfort University. Accessed on March 5, 2002.
URL: www.jtap.ac.uk/reports/htm/jtap-042.html#_Toc454868885

Kahn, Robert, and Robert Wilensky. Framework for Distributed Digital Object Services. May 13, 1995. cnri.dlib/tn95-01. Accessed on March 5, 2002.
URL: www.cnri.reston.va.us/k-w.html

Koehler, Wallace. "Digital Libraries and World Wide Web Sites and Page Persistence". University of Oklahama, 2000. Accessed on March 5, 2002.
URL: http://informationr.net/ir/4-4/paper60.html
This article discusses page persistence and intermittence, defined as the disappearance but reappearance of Web documents. At any given time, about five percent of Web pages are intermittent, which is to say they are gone but will return. Based on a 120-week longitudinal study of a sample of Web documents, it appears that the half-life of a Web page is somewhat less than two years and the half-life of a Web site is somewhat more than two years. That is to say, an unweeded Web document collection created two years ago would contain the same number of URLs, but only half of those URLs point to content.

Miller, James S. "W3C and Digital Libraries". In D-Lib Magazine. November 1996. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/november96/11miller.html

Naming and Linking Strategy. RLG Digital Collections Project. Studies in Scarlet Project. Accessed on March 5, 2002.
URL: http://www.rlg.org/scarlet/name.html

Nelson, Michael L., and B. Danette Allene. "Object Persistence and Availability in Digital Libraries". D-Lib Magazine, January 2002. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/january02/nelson/01nelson.html
Article discusses survey of object persistence and availability of 1,000 digital library (DL) objects. Based on 151 data samples, survey found 3% of the total objects appeared to be no longer available. This matches the value of 3% lost URLs reported by Lawrence et al. In their 2001 study entitled "Persistence of Web References in Scientific Research".

Nelson, Michael L. et al. "Persistence of Web References in Scientific Research". IEEE Computer, 34, (2), 26-31. Accessed on March 5, 2002.
URL: http://csdl.computer.org/comp/mags/co/2001/02/r2026abs.htm

The Open Archives Initiative Protocol for Metadata Harvesting. 2.3. Unique Identifiers - Document Version 2001-04-24. Accessed on March 5, 2002.
URL: www.openarchives.org/OAI/openarchivesprotocol.htm#UniqueIdentifier

Paskin, Norman. "E-Citations: Actionable Identifiers and Scholarly Referencing". Version 1.2, December 17, 1999. Accessed on March 5, 2002.
URL: http://dx.doi.org/10.1000/170

Paskin, Norman. "Information Identifiers". Learned Publishing. Vol 10, no. 2, 135-156. April 1997. Accessed on March 5, 2002.
URL: http://www.elsevier.com/inca/homepage/about/infoident/Menu.shtml

Paskin, Norman. Position paper for W3C Workshop on Digital Rights Management for the Web (22/23 January, 2001) Accessed on March 5, 2002.
URL: www.doi.org/001219_IDF_DRM_pos_paper.htm

Payette, Sandra. "Persistent Identifiers on the Digital Terrain". RLG DigiNews. April 15, 1998. Accessed on March 5, 2002.
URL: www.rlg.org/preserv/diginews/diginews22.html#Identifiers
This article clearly describes URNs and major implementations of persistent identifiers, including PURL, The Handle System and DOI. It also looks briefly at the process of implementing persistent identifiers for digital imaging projects.

Requirements for an Identification Scheme for BIBLINK. May 1977. BIBLINK - LB 4034, Section D4.1 Identification. Accessed on March 5, 2002.
URL: http://hosted.ukoln.ac.uk/biblink/wp2/d2.1/doc0004.htm

Sollins, Karen R. and Jeffrey R. Van Dyke. Linking in a Global Information Architecture. Accessed on March 5, 2002.
URL: http://ana-www.lcs.mit.edu/%20anaweb/links.html

Universal Document Naming Systems. Accessed on March 5, 2002.
URL: http://elib.cs.sfu.ca/USIN/naming.html

"Use of DNS Aliases for Network Services", RFC 2219, 1997. Accessed on March 5, 2002.
URL: www.dns.net/dnsrd/rfc/rfc2219.html
Specifies Internet current best practices.

Web Pages Must Live Forever. Jakob Nielsen's Alertbox. November 29, 1998. Accessed on March 5, 2002.
URL: www.useit.com/alertbox/981129.html

Top of Page

Persistent Identifiers - Specific Systems

The Handle System

A Global System of Handle Servers. Library of Congress, National Digital Library Program (NDLP). Internal documentation. Accessed on March 5, 2002.
URL: http://lcweb2.loc.gov/ammem/award/docs/h-s3.html

Grass, Judith and Arms, Williams Y. A Syntax for Distributed Object Handles. August 3, 1994. Accessed on March 5, 2002.
URL: www.ifla.org/documents/libraries/cataloging/metadata/hdl_form.txt
This document discusses some of the issues involved in designing a format for handles and specifies syntax. This handle syntax specification conforms with the requirements of the DLS and allows object handles to work easily with a large number of existing Internet navigation tools.

"Handle Server Overview". Library of Congress National Digital Library Program Internal Documentation. 1997. Accessed on March 5, 2002.
URL: http://lcweb2.loc.gov/ammem/award/docs/handle-server.html
Overview of Handle System

"Handle System Namespace and Service Definition". IETF Internet Draft. April 2001. Accessed on March 5, 2002.
URL: www.idrm.org/draft-irtf-idrm-handle-system-def-01.txt
This document provides a detailed description of the Handle System namespace, data and service model, as well as its operation and authentication protocol.

More on handles and naming authorities. Library of Congress, NDLP (National Digital Library Program) Internal documentation. Accessed on March 5, 2002.
URL: http://lcweb2.loc.gov/ammem/award/docs/h-s2.html

Sam X. Sun. "Establishing Persistent Identity Using the Handle System". Corporation for National Research Initiatives, 2001. Accessed on March 5, 2002.
URL: www10.org/cdrom/posters/1136.pdf
This article proposes that identity reference used over distributed communication should be defined independent from any attributes and/or the public keys associated with the underlying subject, allowing the identity reference to be persistent over changes made to the attributes and/or the public keys. The papers suggests to separate the trust establishment over distributed communication into three categories: the transport trustworthy, the administrative trustworthy, and the content credential, which simplify trust and trust management. These lead to our proposal to use handles registered under the Handle System as persistent identity reference for Internet communication, and utilize the handle system security service to help trust management over the Internet.

Sun, Sam X et al. "Handle System Protocol (ver 2.0) Specification". IETF Internet Draft. April 2001. Accessed on March 5, 2002.
URL: www.idrm.org/draft-irtf-idrm-handle-system-protocol-01.txt
This document is an Internet draft of Handle system protocol, version 2.

Sun, Sam X et al. Overview of Handles System. August 2000. Accessed on March 5, 2002.
URL: http://www.handle.net/overview-current.html

"Technical Information on Digital Objects, Handles and Repositories". Corporation for National Research Initiatives, 1997. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/february97/cnri/02arms2.html

Top of Page

DOI System

Arms, William Y. "Digital Object Identifiers (DOIs) and Clifford Lynch's Five Questions on Identifiers". October 13, 1997. Accessed on March 5, 2002.
URL: www.arl.org/newsltr/194/arms.html
This article provides background information on DOI in the context of the five (5) questions Clifford Lynch posed about identifier systems in his paper "Identifiers and Their Role In Networked Information Applications."

Atkins, Helen et al. "Reference Linking with DOIs: a Case Study". D-Lib Magazine, February 2000. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/february00/02risher.html

Bide, Mark. "In Search of the Unicorn: The Digital Object Identifier from a User Perspective". February 1998. Accessed on March 5, 2002.
URL: www.bic.org.uk/unicorn2.pdf
Forty-page paper reviews why we need new unique identifiers and looks at early development of the DOI. It looks at specific needs of different users. Useful background for understanding early history of DOI and complexity of different user needs.

"The DOD Does DOI: Massive Project May Put Wheels under DOI Initiative". October 16, 2001. Accessed on March 5, 2002.
URL: www.seyboldreports.com/ebooks/news/
011016-dod.html
Release on Pentagon's decision to convert all of its paper-based records into standardized electronic formats and catalogued using DOIs. While the current three DOI registration agencies have recorded only a few million DOIs, the DOD project promises to expand that number into the trillions.

The DOI Handbook. Version 1.0.0, February 2001. Accessed on March 5, 2002.
NB: Version 3.3, November 2003 available at
URL: www.doi.org/hb.html
This handbook provides detailed information on the DOI, its history and development as well as its components (enumeration, description, resolution, policy development). It also has chapters on role of Maintenance Agency, appointment and role of Registration Agencies and Governance. Appendices include information on DOI syntax, Handle System and INDECS metadata framework, Overview of Application Profiles. Includes practical information on how to obtain a DOI prefix, select a registration agency, become a member of DOI Foundation.

DOI: The Digital Object Identifier System. Developed by the International DOI Foundation. Accessed on March 5, 2002.
www.doi.org/overview/sys_overview_021601.html
Provides a current overview of DOI.

Felici, Jim. "DOI: A Tale of Whoa". The Seybold Report, November 19, 2001. Accessed on March 5, 2002.
URL: www.seyboldreports.com/TSR/subs/0116/
html/news-doi.html (NB: Username and password required)
This article briefly looks at DOI and concludes that no technology comparable to DOI is on the horizon, but publishers have not been that supportive. Support for developing standards through membership fees is running thin. The music and film industry have both come up with their own schemes and W3C and IETF have not shown much interest either. One problem is that it is not strictly a standards group. Another is lack of understanding of DOI and IDF doesn't have the money to get the word out. Giant Chair a major e-book player looked at DOI and rejected it. McGraw-Hill has made

encouraging noises about DOI in its offerings, and the technical and scientific journal world is sold on the concept, but few other publishers are stepping up to bat.

"How the "Digital Object Identifier" Works". Business Week. July 23, 2001. Accessed on March 5, 2002.
URL: www.businessweek.com/magazine/content/01_30/b3742032.htm
General article on DOI.

IDF Terms of Governance and Policies with Registration Agencies, Version 4, April 18, 2001. Access on March 5, 2002.
NB: Version 5, January 2002 available at
URL: www.doi.org/doi-ap/terms.pdf
This 17 page articles details terms of IDF governance and policy with registration agencies and shows the status of policy. Where policy is not agreed upon, recommendation is included.

International DOI Foundation. Annual Review, September, 2000. Accessed on March 5, 2002.
URL: www.doi.org/idf-ann-rpt-10-00.pdf
Thirty-six page report gives review and progress of DOI during 2002. Includes list of strategic alliances and members by class. Includes Treasurers report.

Paskin, Norman. DOI: An Overview of Current Status and Outlook. April 1999. Accessed on March 5, 2002.
URL: www.icsti.org/forum/30/#paskin

The International DOI Foundation. Accessed on March 5, 2002.
URL: www.icsti.org/forum/fo9904.html/#paskin
This paper provides an overview of DOI to April 1999.

Paskin, Norman. "Digital Object Identifier: Implementing a standard digital identifier as the key to effective digital rights management". DOI, April 2000. Accessed on March 5, 2002.
URL: www.doi.org/doi_presentations/aprilpaper.pdf
Paper describes fundamentals of DOI and makes business case.

Paskin, Norman. "Digital Object Identifier: The Case for Supporting the International DOI Foundation", 2001.
URL: Accessed on March 5, 2002. www.doi.org/doi_presentations/010808.pdf
A twenty-page presentation on features and benefits of DOI.

Paskin, Norman. "DOI Commercial Implementations: DOI Update 2001", Slide presentation. 2001. Accessed on March 5, 2002.
URL: www.doi.org/frank_bf_oct01/011010Frankfurt.htm
Sixty-two page slide presentation on DOI with a focus on commercial implementations-CrossRef and Learning Systems.

Paskin, Norman et al. "DOI Implementations", Slide presentation 2001. Accessed on March 5, 2002.
URL: www.doi.org/frank_bf_oct01/011010Frankfurt2.htm
Fourteen page slide presentation focusing on DOI and book publishing and learning objects.

Paskin, Norman. "E-Citations: Action Identifiers and Scholarly Publishing". July 2000. Accessed on March 5, 2002.
URL: www.doi.org/citations4.pdf
This article discusses role of "actionable identifiers" such as DOI in enabling scholarly citations in a digital environment. CrossRef consortium is a practical example being

worked on to-day using DOI.

Paskin, Norman et al. "From One to Many: The Next Stage in the Development of DOI Functionality". A DOI Discussion Paper. August 2000. Accessed on March 5, 2002.
URL: www.doi.org/one2many.pdf
Describes the plans and activities to significantly expand DOI functionality, one area being worked on; the others are implementing metadata structures and the commercial deployment of DOIs via Registration Agencies. DOI is an actionable identifier: clicking on a DOI achieves some action. In the initial implementation, this is a single action: clicking on a DOI takes a user to one URL ('one-to-one'). 'One-to-many' implies one DOI may be used to achieve many possible actions. This is the aim of the full DOI implementation. This paper explores the mechanisms and consequences for the move from 'one-to-one' to 'one to-many'.

Powell, Andy. "Resolving DOI based URNs Using Squid", D-Lib magazine, June 15, 1998. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/june98/06powell.html
This article describes an experimental system that allows DOIs encoded as URNs to be resolved on behalf of Web browsers. It also provides information on the issues associated with long-term identifiers and possible future developments.

"Reference Linking with DOIs: A Case Study". D-Lib Magazine. February 2000. Accessed on March 5, 2002.
URL: www.dlib.org.ar/dlib/february00/02risher.html
This paper discusses the creation of DOI-X, a prototype metadata database designed to support DOI lookups and address the integration of metadata registration and maintenance with basic DOI registration and maintenance, enabling publishers to use a single mechanism and a single quality-assurance process to register both DOIs and their associated metadata. It also contains the lookup mechanisms necessary to access the journal article metadata, both on a single-item lookup basis and on a batch basis, such as would facilitate reference linking. The prototype database was demonstrated in October 1999.

Sidman, David and Tom Davidson. A Practical Guide to Automating the Digital Supply Chain with the Digital Object Identifier (DOI). Content Directions, Inc., 2001. Accessed on March 5, 2002.
URL: www.doi.org/topics/CDI-PRQ2001-article.pdf
This article discussed similarity between the factors affecting adoption of the DOI and the UPC. Greatest value held by both lies in role as a universal product identifier throughout the supply chain, value only materializes when a critical mass of players adopts the standard. DOI promises to follow the same path to success by providing a substantial early, independent return on investment on the revenue side, by making it easier for publishers to get paid by the digital distribution channel, and on the cost side, by limiting the amount of investment necessary for publishers to begin publishing electronically in a more serious way.

Technical Note: DOI and OpenURL. DOI, 2001. Accessed on March 5, 2002.
URL: www.doi.org/topics/010402DOI_OpenURL.pdf
This 4 page paper looks at DOI and OpenURL which are complementary technologies: DOI creates a globally accessible resolvable identifier that can be maintained independently of embedded links; OpenURL provides a resolution service that considers local user information when directing users to a resource.

Top of Page

ARKs

Kunze, J. (2001) The ARK Persistent Identifier Scheme, IEFT working draft, September 7, 2001. Accessed on November 19, 2002.
URL: www.ietf.org/internet-drafts/draft-kunze-ark-04.txt

Kunze, John. (2001). A Metadata Kernel For Electronic Permanence. Journal of Digital Information. 2:2; June 11, 2001. Accessed on March 5, 2002. Available at:
URL: http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Kunze
Permanence of electronic information is tied to persistent identification and to certain kinds of metadata. A series of naming systems have been proposed to support persistent identification, Uniform Resource Name, the Digital Object Identifier, and the Persistent URL, each employing some sort of recognizable character sequence in the identifier to alert the recipient to the name assigning authority's intention to make it persist forever. Archival Resource Key (ARK) is a naming scheme to support long-term durable references to objects, based on the principle that persistence is purely a matter of service. Persistence is neither inherent in an object nor conferred on it by a particular naming syntax or prefix, but is achieved through a provider's successful stewardship of objects and their identifiers. ARK is an identifier that binds an object together with metadata conveying a commitment statement from an identified provider and an object description to furnish complete identification. There are many possible protocols and record formats to service ARKs, but the simple and general format called the Electronic Resource Citation (ERC) meets the demands of realistic permanence services. It starts with the Dublin Core consensus and distills out a subset of four semantic buckets - a metadata kernel - that balances the need for adequate identification of persistent objects and low cost meta generation.

Top of Page

PURLs

Library of Congress. National Digital Library Program. (1997). The Relationship between URNs, Handles and PURLs. August 12, 1997. Accessed on March 5, 2002.
URL: http://lcweb2.loc.gov/ammem/award/docs/PURL-handle.html
This article describes URNs, the Handle System and PURLs, and provides examples of how they work. There is also a brief outline of how handles and PURLs interoperate - PURL resolvers act as proxy handle servers. Conclusions reached are: (1) For an institution committed to maintaining persistent identifiers for the resources they own and manage, running a PURL resolver may be a convenient solution. (2) Whatever form of URN or PURL is employed, need to take responsibility for maintaining their URNs. This is much less of a burden than tracking down and modifying every online link. Both the Handle system and the PURL system incorporate tools for updating records.

OCLC Online Computer Library Center, Inc. (2000). Known problems in the OCLC PURL Resolver software release 2.14. Dublin, OH: OCLC Online Computer Library Center, Inc. Accessed on March 5, 2002.
URL: www.purl.oclc.org/problems.html
Page on problems with Resolver Software 2.1.4.

OCLC Online Computer Library Center, Inc. (2002). PURL. Dublin, OH: OCLC Online Computer Library Center, Inc. Accessed on March 5, 2002.
URL: www.purl.oclc.org/
Home pages of OCLC on PURLs.

Shafer, Keith et al. (1996?) Introduction to PURLS. Dublin, OH: OCLC Online Computer Library Center, Inc. Accessed on March 5, 2002.
URL: http://purl.oclc.org/OCLC/PURL/INET96

Weibel, Stuart et al. (2002). PURLs: Persistent Uniform Resource Locators. OCLC Online Computer Library Center, Inc. Accessed on March 5, 2002.
URL: http://purl.oclc.org/OCLC/PURL/SUMMARY
Quick Summary Page on PURLS

Top of Page

URLS, URNs and URIs

Berners-Lee, T. (1998). Uniform Resource Identifiers (URI) : generic syntax. RFC 2396. Internet Engineering Task Force (IETF). Network Working Group. Accessed on March 5, 2002.
URL: www.ietf.org/rfc/rfc2396.txt?number=2396

Hakala, J. (1999). Uniform Resource Names. Tietolinja News, March 15, 1998. Accessed March 5, 2002.
URL: www.lib.helsinki.fi/tietolinja/0199/urnart.html
An overview of URNs including a basic description of how they work and can accommodate other identifiers; the use of namespaces including how ISBN agencies might use them; the development of a URN resolution service; and the decision of the Conference of European National Librarians (CENL) to implement URNs.

Hakli, Esko and J. Hakala (1998). URN implementation in National Libraries. Helsinki University Library, October 1, 1998. Accessed on March 5, 2002.
URL: http://linnea.helsinki.fi/urn/urnimp.html
Discusses issues relating to URN implementation by national libraries and makes recommendations to the Conference of European National Librarians (CENL). Includes consideration of why national libraries should implement the URN, the possible costs involved, a comparison of URN and DOI, and a list of IETF URN Working group specification documents.

Jones, Dominic. (2001). Avoiding 'Linkrot' on your site. IR Web Report. December 12, 2001. Accessed on March 5, 2002.
URL: www.irwebreport.com/features/010703.htm

Library of Congress. National Digital Library Program. (1997). Examples of mechanisms for keeping URLs relatively persistent. Appendix to: Arms, Caroline R. Access aids and interoperability. (1997). Washington, DC: Library of Congress. National Digital Library Program. Accessed on March 5, 2002.
URL: http://lcweb2.loc.gov/ammem/award/docs/persistence.html
Discusses use of aliases, cgi-bin scripts, logical naming schemes, independent URL resolvers or proxy handle resolver to attain persistence. Gives concrete examples.

Moats, R. (1997). URN syntax. RFC 2141. Internet Engineering Task Force (IETF). Network Working Group, May 1997. Accessed on March 5, 2002.
URL: www.ietf.org/rfc/rfc2141.txt?number=2141

Peacock, Ian. (1998). What is a...URI? Ariadne. 18:December 1, 1998. Accessed on March, 2002.
URL: www.ariadne.ac.uk/issue18/what-is
This article is an introduction to URIs, URLs, URNs and URCs.

Van der Werf-Davelaar, Titia. (1999). Identification, location and versioning of web-resources. URI Discussion paper. Accessed on March 5, 2002.
URL: www.konbib.nl/donor/rapporten/URI.html

Walsh, Norman. (2001). A URN namespace for public identifiers. Cover Pages, May 4, 2001. (draft-urn-publicid-03) . Accessed on March 5, 2002.
URL: http://xml.coverpages.org/walshPINamespaceAnn.html
Describes a URN namespace that is designed to allow Public Identifiers to be expressed in URI syntax. XML external entities have two identifiers: a system identifier and a public identifier. The system identifier is a URI, by definition, but the public identifier is simply a string. Historically, the system identifier of an external entity has been a local, or system-specific identifier while the public identifier has been a more global, persistent name. Public identifiers do not fit neatly into the existing web architecture because they are not legal URIs. Many new specifications (XSLT, XML Schema, etc.) have the implicit or explicit requirement that all external identifiers be URIs. The purpose of this namespace is to allow public identifiers to be encoded in URNs in a reliable, comparable way. This document describes a scheme for representing public identifiers as URNs by introducing a public identifier namespace, 'publicid'.

W3C/IETF URI Planning Interest Group. (2001). URIs, URLs and URNs: clarifications and recommendations 1.0. W3C Note 21: September 2001. Accessed on March 5, 2002.
URL: www.w3.org/TR/2001/NOTE-uri-clarification-20010921

Top of Page

URNs AND NBNs

Hakala, J. (2001). Using National Bibliography Numbers as Uniform Resource Names. RFC 3188. Internet Engineering Task Force (IETF). Accessed on March 5, 2002.
URL: www.ietf.org/internet-drafts/
draft-hakala-nbn-01.txt

Hakala, Juha. URN implementation in national libraries. Accessed on March 5,2002.
URL: http://www.lib.helsinki.fi/urn/urnimp.html

Hakala, Juha. URN Namespace ID Registration for the National Bibliography Number (NBN). Accessed on March 5, 2002.
URL: http://www.lib.helsinki.fi/urn/NBN_registration2.htm Dead link - only reference found http://lists.research.netsol.com/pipermail/urn-n id/1999-August/000014.html

Top of Page

Other Namespaces

Cover, Robin. (2001). Namespaces in XML. The XML Cover Pages, August 23, 2001. Accessed on March 5, 2002.
URL: http://xml.coverpages.org/namespaces.html
XML namespaces "provide a simple method for qualifying element and attribute names used in Extensible Markup Language documents by associating them with namespaces identified by URI references. An XML namespace is a collection of names, identified by a URI reference (RFC2396), which are used in XML documents as element types and attribute names. The 'Namespaces in XML' specification resolves potential name clashes by using the Web addressing infrastructure. The new specification allows authors to mix two or more XML-based languages in one document without conflict or ambiguity, thus promoting the modular development and

reuse of XML languages and applications. The modularity and simplicity of XML technology combined with namespaces paves the way for future developments, such as the work in progress in W3C's XML Schema Working Group, and data exchange based on W3C's Resource Description Framework (RDF) architecture."

Top of Page

Publisher Identification Systems

American Institute of Physics. Publisher Item Identifier as a Means of Document Identification. College Park, MD: American Institute of Physics. Accessed on March 5, 2002.
URL: http://www.aip.org/epub/piius.html

Elsevier Science. (1998). Publisher Item Identifier as a Means of Document Identification. Elsevier Science Ltd. Oxford, UK. Accessed on March 5, 2002.
URL: a href=http://www.elsevier.nl/homepage/about/pii/ NB: Same document as above.

ISSN International Centre. International Standard Serial Number (ISSN). Paris, France. Accessed on March 5, 2002.
URL: www.issn.org

International Standard Technical Report Number (ISRN) ISO 10444:1994. Accessed on March 5, 2002.
URL: www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=18506
NB: Order form and abstract only for priced publication

International Standard Music Number (ISMN). ISO 10957:1993. Accessed on March 5, 2002.
URL: www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=18931
NB: Order form and abstract only for priced publication

International standard book numbering (ISBN) ISO 2108:1992. Accessed on March 5, 2002.
URL: www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=6898
NB: Order form and abstract only for priced publication

National Information Standards Organization (NISO). (1996). Serial Item and Contribution Identifier Standard. ANSI/NISO Z39.56-1996 (Version 2). Accessed on March 5, 2002.
URL: http://sunsite.berkeley.edu/SICI
NB: More authoritative edition available at: www.niso.org/standards/

National Information Standards Organization (NISO). (2002). Book Item and Component Identifier. Draft Standard for Trial Use. Bethesda, MD: National Information Standards Organization. Accessed on March 5, 2002.
URL: www.niso.org/pdfs/BICI-DS.pdf

Publisher Item Identifier Adopted. American Institute of Physics. Accessed on March 5, 2002.
URL: http://www.aip.org/epub/piipr.html

NB: Press release for above citation.

Top of Page

Persistence in Journal Citations

LOCKSS Project Status - Alpha And Beta : Permanent Publishing: Local Control of Content Delivered Via The Web. Accessed on March 5, 2002. URL: http://lockss.stanford.edu/projectstatus.htm

Reich, Vicky and David Rosenthal. (2001). LOCKSS: a permanent web publishing and access system. D-Lib Journal, June 2001. Accessed on March 5, 2002.
URL: www.dlib.org/dlib/june01/reich/06reich.html
Article discusses LOCKSS (Lots Of Copies Keep Stuff Safe), a tool for libraries to ensure their community's continued access to web-published scientific journals.

Top of Page

Australian Resources on Persistence

Dack, Diana. (2001). Persistent Identification Systems. Report on a consultancy conducted by Diana Dack for the National Library of Australia, May 2001. Canberra, Australia: National Library Of Australia. Accessed on March 5, 2002.
URL: www.nla.gov.au/initiatives/persistence/PIcontents.html

Dack, Diana. (2001). Persistence Is A Virtue. A paper presented by Diana Dack, at the Information Online 2001 Conference, Sydney, Australia, 16-18 January 2001. Canberra, Australia: National Library Of Australia. Accessed on March 5, 2002.
URL: www.nla.gov.au/nla/staffpaper/2001/dack.html

National Library Of Australia. (2001). Persistent Identifier Scheme Adopted By The National Library of Australia, September 2001. Canberra, Australia: National Library Of Australia. Accessed on March 5, 2002.
URL: www.nla.gov.au/initiatives/nlapi.html

National Library of Australia. (2000). Request For Quotation No. 00/99, provision of digital object management system for the National Library of Australia. June 2000. Canberra, Australia: National Library Of Australia. Accessed on March 5, 2002.
URL: www.nla.gov.au/dsp/doms/domsrfq.html
Provides details on the functional and technical requirements, selection criteria and responses for NLA's Digital Object Management System.

National Library of Australia. (2000). Safeguarding Australia's Web Resources: Guidelines for Creators and Publishers by National Library of Australia. Canberra, Australia: National Library Of Australia. Accessed on March 5, 2002.
URL: http://nla.gov.au/guidelines/2000/webresources.htm
NB: June 2002 edition available at www.nla.gov.au/guidelines/webresources.html

Top of Page

Organizations

Internet Assigned Numbers Authority (IANA). Domain name services. Accessed on March 5, 2002.
www.iana.org/domain-names.htm
Comprehensive source of information on Domain Name Services.

Internet Corporation for Assigned Names and Numbers (ICANN). Accessed on March 5, 2002.
URL: www.icann.org/

Internet Engineering Task Force (IETF). Accessed on March 5, 2002.
URL: www.ietf.org/overview.html

National Information Standards Organization NISO). (2002). Bethesda, MD: National Information Standards Organization. Accessed on March 5, 2002.
URL: www.niso.org/
Comprehensive source of information on NISO standards.

World Wide Web Consortium (W3C). Accessed on March 5, 2002.
URL: www.w3.org/

Top of Page

OpenURL

OpenURL Demonstrator - Distributed Systems - UKOLN. (2001). Bath, UK: UKOLN. Accessed on March 5, 2002.
URL: www.ukoln.ac.uk/distributed-systems/openurl

National Information Standards Organization (NISO). (2001). NISO Creates OpenURL Standard Committee NISO, March 2001. National Information Standards Organization, Bethesda, MD Accessed on March 5, 2002.
URL: www.niso.org/news/releases/PR-OpenURL.html

Open Name Services. OCLC. Accessed on March 5, 2002.
URL: http://names.oclc.org/openurl

Powell, Andy and Ann Apps. (2001). Encoding OpenURLs in Dublin Core Metadata. Ariadne, Issue 27, March 23, 2001. Accessed on March 5, 2002.
URL: www.ariadne.ac.uk/issue27/metadata

Van de Sompel, Herbert, Patrick Hochstenbach and Oren Beit-Arie. (2000). Open URL syntax description. Accessed on March 5, 2002.
URL: www.exlibrisgroup.com/sfx_openurl_syntax.htm

Top of Page

Miscellaneous

Byrnes, Margaret. (2000). Assigning Permanence Ratings to NLMs Electronic Publications. National Library of Medicine, 2000. Mountain View, CA : RLG. Accessed on March 5, 2002.
URL: www.rlg.org/events/pres-2000/byrnes.html

Byrnes, Margaret. (2000). Defining NLMs Commitment to the Permanence of Electronic Information. ARL : a bimonthly report on research library issues and actions from ARL, CNI, and SPARC. 212:October 2000. Accessed on March 5, 2002. Margaret Byrnes, 2000.
URL: www.arl.org/newsltr/212/nlm.html

Mealing, M. (2002). Dynamic Delegation Discovery System (DDDS), Part One: The Ccomprehensive DDDS Standard. Internet Engineering Task Force (IETF). Accessed on March 5, 2002.
URL: http://search.ietf.org/internet-drafts/draft-ietf-urn-ddds-toc-03.txt
This document specifies the exact documents that make up the complete Dynamic Delegation Discovery System (DDDS) standard. The DDDS is an abstract algorithm for applying dynamically retrieved string transformation rules to an application-unique string. This document along with RFC XXXX, RFC YYYY and RFC ZZZZ obsolete RFC 2168 [8] and RFC 2915 [6] as well as update RFC 2276 [5].

Top of Page

Report - Table of Contents

Created: 1996-05-22
Updated: 2004-07-05

Top of page Important Notices

Localisateurs permanents

Localisateurs permanents pour les publications du gouvernement fédéral canadien

Résumé d'une étude menée pour le compte du Programme des services de dépôt et de la Bibliothèque nationale du Canada

19 novembre 2002

par
Meredith Dickison

Table des matières

SOMMAIRE
Recommandations

SYSTÈMES À CODE D'IDENTIFICATION UNIQUE
A. Aperçu
B. Systèmes

1. Noms de ressources uniformes *(Uniform Resource Names -URNs)*
2. URL permanents (PURLs)
3. Le système Handle *(Handle System)*
4. Code d'identification d'objet numérique *(Digital Object Identifier (DOI) / International DOI Foundation (IDF)*
5. Clé de ressource archivistique *(Archival Resource Key -ARK)*

C. Autres Outils

- LOCKSS (Lots of Copies Keep Stuff Safe)
- LinkOpenly
- OpenURL

D. Autre options

ACCÈS PERMANENT AUX RESSOURCES WEB
Résume des questions a l'étude

- A. Manque de permenance
- B. Coûts et risques
- C. Obstacles à une meilleure gestion des URL
- D. La permanence n'est pas priorité absolue
- E. Le choix du moment pour implanter un système géré de localisateurs

permanents
- F. Besoin immédiat de leadership et de soutien financier

RECOMMANDATIONS

NOTES DE CONCLUSION

BIBLIOGRAPHIE

Haut le page

Sommaire

Effectuée au cours des mois de février et de mars 2002, l'étude avait pour but d'examiner les recherches et les pratiques actuelles dans le domaine des localisateurs et des codes d'identification d'objets numériques permanents et de recommander une suite de mesures pour mettre en œuvre un système de localisateurs de ressources uniformes permanents pour les publications du gouvernement fédéral. Durant la période des entrevues, l'auteure a effectué une étude et une analyse documentaires complètes sur la recherche existante. Elle a interrogé - en personne ou par correspondance - des employés du gouvernement fédéral responsables des sites Web du gouvernement fédéral et correspondu avec des chercheurs œuvrant dans des projets relatifs aux localisateurs permanents. Les personnes interrogées ont ensuite examiné le rapport préliminaire, formulé des recommandations ou des améliorations, puis approuvé la version finale du rapport ainsi que ses recommandations.

Aujourd'hui, le gouvernement a établi un cadre de gestion de l'information électronique durant tout son cycle de vie et produit des normes et des lignes directrices concernant la publication de ressources électroniques sur le World Wide Web (le Web), qui forment une assise pour la promotion de l'accessibilité et de la permanence des URL. Cependant, étant donné que les localisateurs de ressources uniformes (URL) - codes d'identification uniques permettant la consultation de ressources électroniques sur le Web - sont liés à l'emplacement, tous les liens menant à une ressource donnée sont brisés lors de son déplacement ou de son élimination, interdisant ainsi l'accès à la ressource. Bien que les estimations divergent, des sources indiquent que la durée de vie d'une publication électronique type varie entre 44 jours et deux (2) ans.

Après avoir mené une série d'entrevues dans différents ministères, l'auteure conclut qu'individuellement, certains ministères appliquent des " pratiques exemplaires " dans leur gestion des URL (dont le réacheminement), favorisant ainsi la permanence, ainsi que des approches holistiques de gestion des ressources électroniques. Cependant, il semble que le fondement de cette démarche doive être renforcé. Les employés semblent éprouver de la difficulté à appliquer concrètement certaines lignes directrices et la mise en œuvre n'est pas uniforme d'un ministère à l'autre. En outre, certaines lignes directrices ne sont pas suffisamment rigoureuses ni complètes. Enfin, des employés semblent ne pas connaître l'ensemble des normes et des lignes directrices concernant la gestion de l'information électronique.

La Bibliothèque nationale du Canada et le Programme des services de dépôt, Communications Canada, possèdent des compétences particulières en ce qui concerne la classification et la gestion de l'information. Ces compétences devraient être mises à profit dans le domaine des publications électroniques du gouvernement fédéral, afin d'assurer la permanence et l'accessibilité à long terme de sources d'information fédérales importantes. La question de la permanence ne concerne pas qu'une seule institution mais bien l'ensemble du gouvernement, de sorte que les solutions doivent être élaborées par voie de collaboration, dans l'ensemble du

gouvernement. Le gouvernement fdééral s'est engagé à mettre en œuvre l'initiative Gouvernement en direct (GED) d'ici 2005. Il est donc crucial de chercher immédiatement une solution à la question de l'accessibilité. On doit fournir dès maintenant le soutien financier nécessaire pour mettre en œuvre des solutions efficaces à la question de la permanence.

Recommandations :

- La Bibliothèque nationale du Canada et le Programme des services de dépôt doivent examiner la question de la permanence et créer un groupe de travail ou une commission pour en faire un agent de promotion de la permanence. Le groupe de travail doit obtenir le droit de siéger au Comité consultatif du Secrétariat du Conseil du Trésor sur Internet afin d'assurer sa présence auprès des principaux décideurs et d'établir une voie lui permettant de communiquer ses connaissances spécialisées sur la question.

- Le groupe de travail devrait réunir des gestionnaires et des membres du personnel de différents ministères et organismes du gouvernement fédéral, œuvrant dans des domaines variés et aptes à présenter une perspective différente.

- Le groupe de travail devrait s'efforcer d'exercer une influence et de mettre à profit ses connaissances dans le but de promouvoir la permanence dans la gestion des ressources d'information électroniques du gouvernement fédéral. À cette fin, il doit mener des activités de sensibilisation et de promotion, renforcer les normes et lignes directrices fédérales et rechercher des appuis financiers pour les efforts de collaboration et la participation à des projets pilotes.

  L'inaccessibilité de l'information associée à la non-permanence des adresses URL pose un risque très élevé. Chaque jour, des utilisateurs publics ont besoin de l'information gouvernementale. On ne peut chiffrer en termes financiers le temps perdu à chercher l'information ni les pertes éventuelles associées au fait de ne pouvoir la trouver ni la consulter. Le gouvernement élabore des passerelles Web toujours plus vastes et complexes et tend à offrir un nombre croissant de services en direct. Dans ce contexte, la permanence des liens vers l'information électronique et l'accessibilité des ressources sont des facteurs encore plus cruciaux.

Haut le page

Systèmes à code d'identification unique

A. Aperçu

Essentiellement, il existe cinq systèmes de code d'identification officiels, qui font tous l'objet d'une documentation abondante.

1. Noms de ressources uniformes / codes d'identification de ressources uniformes *(Uniform Resource Names / Uniform Resource Identifiers - URN/URI)*
2. Localisateurs de ressources uniformes permanents *(Persistent Uniform Resource Locators - PURLs)*
3. Système Handles *(Handle System)*
4. Code d'identification d'objet numérique *(Digital Object Identifier - DOI)*
5. Clé de ressource archivistique *(Archival Resource Key - ARK)*

Ce sont tous des systèmes " gérés " s'appuyant sur une structure de registre qui nécessite une maintenance afin de garantir que les codes d'identification attribués demeurent uniques et permanents et qu'ils seront mis à jour en fonction des changements de propriété ou autres.

Les limites des URL ont été constatées dés le départ, ce qui explique que les travaux visant à élaborer un système de codes d'identification uniques soient en cours depuis 10 ans. Malgré cela, les systèmes examinés ont tous leurs forces et leurs faiblesses et en sont essentiellement à leurs dbéuts, à l'exception des PURL. Dans tous les systèmes, on note une indépendance entre le système de " dénomination " et le système de résolution, ce qui, bien sûr, fait contraste avec les URL, qui portent leurs propres données de résolution, comme le système hôte à contacter et le protocole à utiliser. Tous les systèmes de dénomination sont déployés aujourd'hui, mais aucun ne l'est à grande échelle. Cela s'explique par la lente évolution du processus d'élaboration et de déploiement de systèmes de résolution utilisables par les fureteurs existants.

Tous ces systèmes de dénomination sont de type " parapluie ". Ils peuvent donc englober des systèmes d'identification additionnels, de sorte que rien n'interdit qu'ils soient utilisés en conjonction avec d'autres systèmes éditeurs de numérotation ou d'identification existants.

Le principal avantage qu'offrent ces systèmes (à l'exception des PURL), outre celui d'établir un code d'identification unique, est la possibilité de résolution multiple. La résolution la plus simple est la résolution un à un comme, par exemple, un URN vers un URL ou un autre URN. Selon le système hiérarchique des résolveurs, dans lequel chaque résolveur fournit l'information nécessaire pour porter le processus au prochain niveau, il est aussi possible de transformer un URN (ou un DOI, un Handle ou un ARK) en URL multiples associés à l'emplacement de la ressource, à l'adresse d'un répertoire contenant la ressource, à une description bibliographique de la ressource ou à un formulaire de commande pour obtenir la ressource, etc. Ces systèmes offrent de nouvelles possibilités, mais leur émise en œuvre concrète pose des difficultés additionnelles comme, par exemple :

1. granularité : quelles sont les parties des documents nécessitant une identification et un accès distincts ?

2. contrôle des versions : à partir de quel point un document a-t-il changé suffisamment pour justifier une identification distincte ?

3. droit d'auteur : les parties de documents assujetties à des droits d'auteur différents doivent-elles porter une identification distincte ?

Les données ou métadonnées additionnelles requises pour l'identification et la localisation des ressources peuvent être intégrées à la chaîne d'identification, à la chaîne de localisation ou au système de résolution. Bien qu'ils puissent varier d'un système à l'autre, les groupes des métadonnées utilisés appliquent généralement le principe permettant l'utilisation d'un ensemble de données restreint - un noyau (dans le système ARK, cela signifie aussi peu que quatre éléments).[1]

Haut le page

B. Systèmes 1. Noms de ressources uniformes *(Uniform Resource Names - URNs)*

Les URN sont un système de dénomination comprenant un code d'identification d'espace de noms (Namespace Identifier Code) inscrit auprès de l'IANA (Internet

Assigned Numbers Authority) et une chaîne propre à l'espace de noms (Namespace Specific String), code unique et permanent servant à identifier le document. Il n'existe pas présentement de système de résolution uniformisé pour les URN, bien que l'Internet Engineering Task Force (IEFT) ait produit une proposition concernant un résolveur appelé le Dynamic Delegation Discovery System.[2] Les URN font partie de l'architecture URI (Uniform Resource Identifier - code d'identification de ressources uniformes), qui se veut un service de recherche de ressources complet et permanent.

Quelques URN ont été inscrits et mis en œuvre (OASIS, [3] par exemple) mais, en l'absence de résolveur uniformisé, ils utilisent des techniques comme des catalogues qui assignent des URN à des URL pour fonctionner, ou encore des serveurs mandataires ou des résolveurs expérimentaux.

Avantages

- Offre un système de dénomination souple.
- Non lié à un protocole ou à un emplacement particuliers.
- Permet que l'URN réfère à de multiples occurrences d'un article.
- Permet le déploiement dans un modèle centralisé ou décentralisé.
- Permet l'adaptation à tout modèle d'identification présentement en usage.
- Fait la distinction entre les modèles URL et les modèles à adresses et suppose une volonté de permanence chez l'auteur de l'attribution.
- S'inscrit dans les travaux de l'IEFT / W3C.

Inconvénients

- On ne note pas de déploiement significatif pour le moment ni de service de résolution uniformisé pour les codes d'identification des URN.
- Les perspectives de déploiement sont incertaines ; W3C travaille à cette question, mais sans grande vigueur.[4]
- Le système n'est pas accessible présentement par les fureteurs habituels, sauf par le biais d'un serveur mandataire (de sorte que les codes d'identification doivent être encapsulés dans l'URL au moyen de l'adresse du serveur mandataire).
- Aucun module d'extension de fureteur n'est disponible.
- Différents supports de document utilisent des URN, mais des questions demeurent en suspens quant à l'utilisation en XML.
- Il n'y a pas de système de soutien : il n'existe pas présentement d'outils d'administration et de gestion.

La Conférence des directeurs de bibliothèques nationales a donné son accord de principe à l'utilisation des URN en parallèle avec le NBN (National Book Number - Numéro de livre national). On n'a pas établi de système de résolution, mais cette fonction s'effectuera probablement par le biais de bases de données bibliographiques nationales ou d'un index ou d'archives Web. Les pays nordiques utilisent cette application URN, la résolution relevant d'une application *Nordic Web Index*.[5]

Haut le page

---

2. URL permanents (PURLs) *(Permanent URLs)*

Les URL permanents ont ceci de particulier que presque que toutes les sources

documentaires y voient une solution " provisoire " à la question de la permanence. Cette règle fait toutefois exception :

> On croit généralement que le Service de résolution des PURL précède le système de résolution à URN et que les PURL seront périmés dès que toutes les normes URI seront mises en œuvre. Cependant, un système PURL est un système de résolution URI qui permet une résolution simple un à un d'un PURL à un URL. Ce type de résolution simple à des fins de gestion des URL va demeurer important pour certains types de collections de ressources Web, aujourd'hui et plus tard. Il est très probable que des systèmes de résolution URN complexes vont être crées pour contrôler l'accès aux ressources de collections particulières (commerciales). Il est même probable que ces systèmes complexes seront plus faciles à gérer si la location-résolution s'effectue par un système de type PURL, faisant l'objet d'une maintenance distincte. Les exigences fonctionnelles de la gestion des URL et celles liées à d'autres paramètres d'accès peuvent justifier la mise en œuvre de systèmes distincts.

À ce sujet, voir le document " Identification, Location and Versioning of Web Resources ".

Un PURL fonctionne comme un URL, à la différence qu'il renvoie à un service de résolution intermédiaire plutôt que de mener directement à l'emplacement de la ressource Internet. Un service de résolution PURL associe le PURL et l'URL, qu'il retourne au client. Celui-ci peut ensuite terminer l'opération de la façon normale. Les PURL utilisent les protocoles courants HTTP pour se brancher aux résolveurs PURL ainsi que des réacheminements standard HTTP pour renvoyer l'information à l'auteur de la demande. Quelqu'un doit se charger d'opérer les résolveurs PURL qui assurent cette permanence. Si l'URL associé au PURL devient périmé, la résolution du PURL peut faire défaut. Le PURL et son registre historique complet demeureront disponibles tant que le service PURL lui-même est maintenu.

Online Computer Library Centrer Inc. (OCLC) offre le logiciel PURL Resolver sur le Web. L'OCLC ne consigne pas le nombre d'institutions qui utilisent son logiciel ; toutefois, les PURL sont assez largement déployés. La National Library of Australia utilise les PURL, sans en être totalement satisfaite. La NLA estime que ce système n'a pas été largement adopté, tout en reconnaissant que cela peut être attribuable au fait que le système n'a pas joui d'une publicité importante.[6]

Il semble que le U.S. Government Printing Office (GPO) dispose de la plus importante installation PURL au monde. étant donné que le GPO offre aux bibliothèques de dépôt fédérales des états-Unis l'accès aux documents électroniques par le biais de GPO Access, on suppose qu'il est responsable de leur permanence, malgré qu'aucune exigence juridique ne prescrive explicitement l'obligation d'assurer la permanence des versions électroniques en ligne des documents gouvernementaux. Toutefois, le GPO estime qu'il est juridiquement tenu d'assurer l'accès permanent à l'information électronique.[7]

Avantages

- Constitue une solution simple utilisant des réacheminements standard HTTP et des URL ; il ne nécessite donc pas une formation poussée.
- Comprend un ensemble de logiciels utilitaires facilitant la gestion des URL en permettant à l'usager de mettre à jour ses propres ressources, ce qui constitue un aspect crucial de l'intégrité et du maintien des liens.
- On peut déjà se procurer un logiciel d'OCLC.
- Nombre d'organisations pourraient déployer un serveur PURL, de sorte qu'il serait possible de déployer un modèle centralisé ou décentralisé.
- L'utilisation d'un PURL témoigne de la volonté d'assurer la permanence de la ressource.

Inconvénients

- Ce n'est pas une norme, seulement une technique.
- Lorsque l'éditeur interrompt l'accès aux services, aucun suivi ne peut être assuré.
- Les descriptionS d'objet ou d'énoncés d'engagement ne sont pas nécessaires.
- Ne peut être dissocié de l'emplacement.
- Dépend du protocole.
- Ne permet qu'une résolution un à un.
- Les fureteurs habituels retournent l'URL dans le champ de localisation. (Cela deviendrait la référence dans bien des cas, à moins que le PURL soit clairement indiqué dans ou sur l'objet même ou que les fureteurs soient modifiés de façon à afficher l'URL permanent demandé plutôt que l'URL réel représentant l'emplacement actuel de la ressource.)
- Presque toujours vu comme une solution provisoire et non une réponse définitive à la question de la permanence.
- Le logiciel pose quelques difficultés.[8]

Haut le page

---

3. Le système Handle *(Handle System)*

Handle est un système de dénomination permanent réparti élaboré par la Corporation for National Reserach Initiatives (CNRI). Les " pointeurs " (handles) sont des noms uniques d'objets numériques, soutenus par une fonction de résolution consistant en un réseau réparti de serveurs mondiaux, régionaux et à antémémoire. Un Registre mondial des pointeurs (Global Handle Registry), exploité par la CNRI, répertorie les autorités de dénomination de premier niveau.

Le système Handles a été déployé à la Library of Congress Digital Library[9] avec deux serveurs (un serveur de pointeurs et un " serveur de pointeurs mandataire ", des pointeurs étant ajoutés aux entrées MARC et aux repères EAD) ainsi qu'au U.S. Copyright Office.[10] Handles a aussi été mis en œuvre à D-Space, projet conjoint de la MIT Library et de Hewlett-Packard regroupant 10 000 articles produits chaque année par des membres du corps professoral et des chercheurs du Massachusetts Institute of Technology.[11] Le système a aussi été mis en œuvre par l'International DOI Foundation, décrite à la page suivante.

Selon les renseignements de la CNRI, le système Handles n'a été déployé qu'à un seul endroit au Canada.[12]

Avantages

- Ce système a fait ses preuves et est déjà déployé dans des applications de premier ordre. Il sera interopérable avec le DOI, si ce système en vient à être largement développé.
- Le système de dénomination est souple et compatible avec les URN.
- Comporte un mécanisme de résolution.
- Peut être déployé dans un modèle centralisé ou décentralisé.
- Ne dépend pas d'un protocole en particulier.
- Permet une résolution multiple.
- Le logiciel et les outils de gestion sont fournis par la CNRI.
- élaboré spécifiquement pour les applications de bibliothèques numériques.

- Peut s'adapter facilement aux modèles existants.
- Utilisable en tandem avec LDAP ou WHOIS++, dans le contexte d'un service de recherche inverse.

Inconvénients

- Sa mise en œuvre est complexe. [13]
- Ne peut pas encore être utilisé avec les fureteurs conventionnels sans l'ajout d'un module d'extension pour le client.
- N'est pas enregistré comme modèle URI / URN (la marche à suivre pour ce faire ainsi que les avantages relatifs de l'enregistrement en tant qu'URI ou URN ne sont pas clairs). Il est peu probable qu'il soit mis en œuvre par les développeurs de fureteurs conventionnels sans la reconnaissance formelle de W3C.
- En tant que projet de recherche, la CNRI pose la question du soutien du système si le financement est retiré et que l'organisme cesse d'exister (bien qu'étant donné l'investissement de la LC et de l'IDF, il est peu probable que le système s'effondre soudainement).[14]
- Le système comporte certaines limites en matière de sécurité.[15]
- La résolution d'infrastructure de droits exclusifs complexes nécessite un tableau de consultation directe.
- Les descriptions d'objet ou d'énoncés d'engagement ne sont pas nécessaires.

Haut le page

---

4. Code d'identification d'objet numérique *(Digital Object Identifier (DOI) / International DOI Foundation - IDF)*

Ce système est géré par l'International Digital Object Identifier Foundation et comporte trois éléments :

1. code d'identification permanent de la propriété intellectuelle, assorti de règles de syntaxe et d'attribution
2. système de résolution utilisant la technologie du système Handles
3. structure de métadonnées identifiant les articles représentés par le DOI (en fonction de l'ensemble de métadonnées de l'INDECS)

Les DOI sont attribués par les autorités d'enregistrement (il en existe présentement quatre[16] : trois aux états-Unis et une en Corée) et enregistrés dans le registre mondial.

Le modèle commercial n'est pas encore fixé définitivement,[17] : mais l'IDF est conçu comme un organisme fondé sur le principe du recouvrement des coûts. Les questions de gestion ne sont pas encore réglées.

Bien que 200 organisations [18] : attribuent des DOI et que plus de 4 millions de DOI aient été créés en date de janvier 2002, les sources documentaires ne sont pas unanimes au sujet des DOI. Le système a été adopté par le U.S. Department of Defense,[19] : ce qui joue en sa faveur. D'autres estiment que le système n'a pas fait ses preuves et qu'il pourrait faire long feu, bien qu'on n'entrevoie encore aucune technologie qui y soit comparable.[20] : Malgré qu'il ait été conçu pour répondre aux besoins particuliers du domaine de l'édition, le système n'a pas été adopté par tous les éditeurs et il est en voie de s'adapter aux besoins d'autres utilisateurs.[21] :

Au Canada, les Presses scientifiques du CNRC assignent des DOI aux articles de

revues.[22] : Les Presses scientifiques paient une cotisation de 750 $ US et 0,60 $ par DOI pour le dépôt et 0,10 $ pour la recherche, bien que les coûts de recherche puissent être éliminés par un système d'antémémoire local. Les Presses publient les DOI dans les versions imprimées de leurs revues. Leur site Web donne des indications sur la façon d'établir des liens vers les articles au moyen des DOI (par exemple, en utilisant CrossRef) et de liens URL (pseudonymes) afin d'assurer un lien stable et à long terme vers les différentes pages d'accueil des revues. Les Presses scientifiques sont satisfaites de leur expérience des DOI à ce jour et trouvent que le système est facile à maintenir (essentiellement, on envoie chaque mois un fichier contenant les numéros de DOI et les métadonnées assignés à CrossRef). Les Presses scientifiques estiment que les DOI sont un des moyens permettant d'assurer la permanence des ressources, mais elles les utilisent aussi pour obtenir une plus grande visibilité sur le marché.

Avantages

- C'est un système entièrement géré applicable dans un modèle centralisé ou décentralisé.
- Son système de dénomination souple peut s'adapter rapidement aux modèles existants.
- Comprend une infrastructure de résolution.
- Ne dépend pas d'un protocole en particulier.
- Permet une résolution multiple.[23] :
- Le système est soutenu par des logiciels offerts gratuitement (nombre d'avantages communs avec le système Handles).
- Au minimum, le système peut être largement déployé par de grands éditeurs commerciaux.
- Les institutions ont la possibilité de devenir autorités d'enregistrement, ce qui leur permet de créer des services à valeur ajoutée.
- Le système peut être développé dans un environnement de normes ouvertes.
- IDF a entrepris l'intégration à des activités liées à des normes connexes pour JPEG (participation au projet Rights Data Dictionary - projet de dictionnaire des données sur les droits).
- IDF commence à obtenir un appui d'entreprises technologiques comme Adobe Acrobat et Microsoft.[24] :

Inconvénients

- Le système pose un risque élevé en ce qu'il n'a pas fait ses preuves. On ignore s'il sera largement déployé à long terme.
- Le passage à un système à gestion intégrale est coûteux. En ce moment, de nombreux facteurs de coût demeurent incertains.
- Bien des questions se posent sur la mesure dans laquelle les intérêts des grands éditeurs vont prévaloir. À long terme, les coûts pourraient être inabordables pour les petits éditeurs et les propriétaires de sites Web. On doit se demander si les grands éditeurs ne vont pas restreindre l'accès au système plus tard.
- Bien des incertitudes persistent quant aux questions de gestion associées à l'établissement d'un organisme d'enregistrement des DOI (les règles actuelles concernant l'enregistrement sont plutôt vagues), si bien que l'évolution du déploiement demeure incertaine.
- Pour le moment, le système ne peut pas encore être utilisé avec les fureteurs conventionnels sans l'ajout d'un module d'extension pour le client (bien qu'on puisse utiliser des serveurs mandataires).
- Devenir organisme d'enregistrement serait coûteux et la valeur d'une telle décision serait discutable en l'absence d'une stratégie de commerce électronique.
- IDF a montré peu d'intérêt à l'égard des bibliothèques et archives et de leurs

besoins.

- Mandat d'utilisation du système Handle pour la résolution DOI.
- Enregistrement sur la base de services tarifiés.
- Les descriptions d'objet ou d'énoncés d'engagement ne sont pas nécessaires à l'exception d'une partie du système CrossRef basé sur des services tarifiés.

Haut le page

---

5. Clé de ressource archivistique *(Archive Resource Key - ARK)*

" Un des fondements du ARK est que la permanences est purement un rôle de soutien. " - John A. Kunze

État conçu pour les gardiens des objets numériques archivés, le système ARK comprend une syntaxe des codes d'identification et trois services offrant à l'utilisateur un lien entre un objet et une promesse de gestion de cet objet, ainsi qu'une description de l'objet et un lien ver ce dernier. La citation de ressource électronique (ERC), une méthode de prestation des services ARK qui est en cours d'élaboration, consiste d'un format de métadonnées constitué de quatre éléments dérivés de Dublin Core. L'ERC fournit un moyen ultra simple et compact de créer et de traiter des descriptions d'objet.[25] En comparaison de XML, l'ERC est peu efficace lorsque vient le temps d'imbriquer es dossiers entièrement transparents à une profondeur dépassant quelques niveaux.

L'expérience de mise en œuvre du système se limite au service prototype ARK établi à la U.S. National Library of Medicine.[26] Un développeur du système l'avoue : " La prestation des services ARK nécessite des conditions très particulières." [27]

Avantages

- Système conçu spécialement pour les documents d'archive numériques.
- Indépendant de l'emplacement et du protocole.
- Utilisable avec les fureteurs conventionnels d'aujourd'hui (ne requiert qu'une modification mineure des fureteurs pour leur permettre de remplacer la chaîne " http " par une chaîne " ark " aux fins de résolution).
- Possibilité de résolution multiple.
- Traite directement les problèmes de service après la permanance.
- Distingue clairement le rôle du nom de la tâche de départ du rôle de soutien de la tâche en cours pour l'accès et la description.
- Recherche de solution, lorsque nécessaire. Peut être mis en application avec un petit caractère scripte qui utilise le système de nommage de domaine Internet pour localiser l'actuel fournisseur de service.
- Services associés utilisant les métadonnées (possibilité de recycler les métadonnées existantes).

Inconvénients

- Système trop récent pour en évaluer les perspectives d'acceptation : peut être jugé trop contraignant par certains.
- Comme le URN et le DOI, nécessitera un jour une infrastructure de mécanisme de résolution de problème - Il n'y a aucune garantie que ceci sera entièrement élaboré ou accepté.
- Trop tôt pour évaluer de façon réaliste les coûts de mise en œuvre d'un tel

système, mais il est clair que cela nécessitera un investissement considérable.

Haut le page

C. Autres Outils

LOCKSS (Lots of Copies Keep Stuff Safe)
Les sources documentaires couvrent un autre outil d'identification appelé LOCKSS, défini comme un système permanent de publication et d'accès Web.[28]

LOCKSS est un outil permettant aux bibliothèques d'assurer à leur collectivité un accès continu aux revues scientifiques publiées sur le Web. Les bibliothèques peuvent utiliser des mémoires cache Web pour différentes revues. Le mémoires cache recueillent le contenu à mesure qu'il est publié et ne sont jamais vidées. Elles fonctionnement dans un réseau de point à point dans lequel elles détectent et réparent les pages endommagées ou manquantes. Les mémoires cache fonctionnent sur du matériel PC générique utilisant un logiciel d'exploitation libre et ne nécessitent pratiquement pas d'administration spécialisée, ce qui leur permet de conserver une revue à un coût abordable. Mis à l'essai en 2000 dans plus de 40 bibliothèques de par le monde et avec la collaboration de plus de 30 éditeurs, le système s'est avéré techniquement utilisable. D'autres essais sont présentement en cours (fin des travaux prévue pour août 2002),[29] : mais les perspectives de déploiement sont inconnues pour le moment.

Haut le page

LinkOpenly
Un autre logiciel d'exploitation libre appelé LinkOpenly[30] : est en voie d'élaboration par Openly Informatics Inc., de concert avec le système DOI de CrossRef. Ce logiciel vise à fournir un lien permanent et complet vers les revues électroniques à partir de pages Web, de bibliothèques et de bases de données. Openly Informatics Inc. a déjà créé OpenlyJake,[31] : qui propose un autre mécanisme de liaison par la gestion de métadonnées de ressources en ligne au moyen d'un répertoire collectif de bases de données, d'un contrôle de l'autorité des titres et d'outils de liaison, en plus de faciliter l'ajout d'une couche locale de fonds documentaires.[32]

Haut le page

OpenURL
Au Royaume-Uni, le milieu des bibliothèques mène aussi des travaux dans ce domaine. OpenURL[33] : fournit un mécanisme permettant d'encoder en URL une référence relative à une ressource d'information, généralement une ressource bibliographique. L'OpenURL est, en fait, un URL apte à exécuter une action et qui transporte des métadonnées ou des clés donnant accès à des métadonnées pour l'objet associé à l'OpenURL. L'OpenURL cible un résolveur OpenURL offrant des services localisés dans un environnement de liens ouverts. En général, on réfère à l'OpenURL en tant que composante du service institutionnel (Institutional Service Component - ISC) de l'utilisateur. Le reste de l'OpenURL transporte la référence. Celle-ci est fournie par l'application d'un code d'identification pour la ressource (par

exemple, un code d'identification d'objet numérique) ou l'encodage de métadonnées sur la ressource comme, par exemple, le titre, l'auteur, le titre de la revue, etc., ou par une combinaison de ces deux moyens. Il est aussi possible d'encoder un identificateur local pour la ressource à l'intérieur de l'OpenURL. Associé à l'information sur le lieu de création de l'OpenURL, ce mécanisme permet au logiciel qui reçoit l'OpenURL de demander des métadonnées additionnelles au sujet de la ressource d'information.

Le Committee AX, comité des normes de la National Information Standards Organization (NISO), a adopté récemment une version provisoire d'OpenURL en tant que Version 0.1 de la norme. Le Comité reconnaît que la technologie OpenURL en est à ses débuts, mais il est persuadé que ce système n'a fait qu'effleurer toutes les possibilités qu'il peut offrir en matière de référence bibliographique. Le Comité prévoit que, dans un avenir rapproché, des tierces parties pourraient fournir des liens OpenURL pour n'importe quelle ressource d'information et que des liens OpenURL pourront être encodés en XML.[34]

Haut le page

D. Autres options

Outre l'option d'adopter des systèmes de codes d'identification existants ou d'élaborer et de gérer son propre système (approche choisie par la National Library of Australia), il serait possible de promouvoir la permanence des ressources en utilisant RDF/XML comme cadre commun de description des ressources ou base d'encodage et d'échange de données.[35] : À partir des éléments de métadonnées du Dublin Core déjà requis pour les ressources Web gouvernementales, il pourrait être possible d'inclure un code d'identification qui pourrait permettre l'identification unique des publications électroniques. D'autres travaux devront être menés pour établir le mode de fonctionnement de ce système et le délai nécessaire à sa réalisation ainsi que la place qu'il occuperait dans le contexte des systèmes d'identification existants (les GIL, par exemple).[36]

Haut le page

Accès permanent aux ressources Web

Une autre question fait l'objet de débats et de discussions dans les sources documentaires sur la permanence : la relation entre la permanence de l'identification et la permanence de l'accès. Les codes d'identification permanents dont nous avons discuté fournissent un moyen d'identification permanente et d'accès permanent tant que l'article est en ligne et que le système est maintenu. Cependant, les institutions souhaitent-elles vraiment " garantir " que tous les articles possédant un code d'identification permanent vont demeurer indéfiniment en ligne ?

Un des éléments cruciaux de la question de la permanence consiste à établir une définition précise permettant à une institution de communiquer clairement son degré de détermination à maintenir l'accessibilité à une ressource en ligne. (Quelques institutions ont affiché en ligne une politique de la permanence. Voir, par exemple, le site de W3C.)

Il semble que la U.S. National Library of Medicine ait effectué le travail le plus concret à ce chapitre en établissant trois catégories principales de permanence des ressources

électroniques, bien qu'on n'ait pas encore regroupé les ressources[37] : dans ces catégories et qu'on ait pu en élaborer une version condensée dans une réitération subséquente.[38] : Les catégories établies par la U.S. National Library of Medicine sont les suivantes :

1. **Validité de l'identificateur**, définie comme la mesure dans laquelle un utilisateur peut avoir l'assurance qu'un identificateur donné ne changera pas. Cette catégorie comporte deux options :

   i) **indéfini ou transitoire** - qui signifie qu'aucune classification n'est entreprise ou que l'identificateur pourrait être changé ou réaffecté, et

   ii) **garanti** - qui signifie que l'identificateur ne sera pas changé ni réaffecté.

2. **Disponibilité de la ressource**, définie comme la mesure dans laquelle une ressource donnée demeurera accessible. Cette catégorie comporte deux options :

   i) **sans garantie** - la ressource peut devenir inaccessible sans préavis, et

   ii) **disponible en permanence** - qui signifie que l'accessibilité de la ressource est garantie, ce qui implique la nécessité de tenir un système d'archives.

3. **Invariabilité du contenu**, ou mesure dans laquelle le contenu d'une ressource donnée et les liens qu'elle contient demeureront inchangés. Cette catégorie comporte quatre options :

   i) **non classifié** - qui signifie qu'on n'a pas entrepris de classification ni donné de garantie,

   ii) **dynamique** - qui signifie que le contenu peut être remplacé, corrigé et révisé et que les liens pourraient changer,

   iii) **statique** - le contenu ne peut faire l'objet que de corrections ou d'ajouts mineurs et les liens internes seront mis à jour, et

   iv) contenu statique - qui signifie que le contenu ne changera pas.

Dans bien des cas, il est possible d'assigner une classification aux catégories de ressources (par exemple, annonces, rapports annuels), de sorte qu'il ne serait pas nécessaire de classifier les ressources individuelles.

La National Library of Australia a aussi effectué des travaux dans ce domaine. Selon la NLA, on peut généralement regrouper sous deux types l'information qui doit être gérée pour assurer l'accès permanent :

a) **ressources statiques** - éléments de version définitive, qui ne seront pas modifiés;

b) **ressources dynamiques** - documents qui subiront des modifications, soit une augmentation par ajout de contenu ou des changements par modification, élimination ou insertion de contenu.

Les ressources peuvent être de type " **archivistique** " - documents qu'on souhaite conserver et auxquels on veut permettre un accès permanent - ou de type " **éphémère** " - documents qui ne seront vraisemblablement pas cités dans d'autres ressources mais qui, souvent, seront marqués d'un signet par les utilisateurs qui recherchent la plus récente version du document. Cette catégorie ne nécessite la permanence de référence que pour la version actuelle.

Dans ce modèle, la décision sur le degré de permanence d'accès à offrir sera affaire de jugement et devra être prise au moment de la création du document.[39] :

Haut le page

Résumé des questions à l'étude

A. Manque de permanence

- Les URL ne sont pas permanents et changent pour toutes sortes de raisons - restructuration des ministères, réorganisation des structures de dossiers par les employés, retrait de pages Web, etc. -, ce qui a pour effet de rompre les liens. Cette non-permanence peut se produire dans n'importe quel ministère ; c'est donc une question qui intéresse l'ensemble du gouvernement et non pas seulement la Bibliothèque nationale et le Programme des services de dépôt. On pourrait documenter ce problème, étant donné que les ministères vérifient les liens et tiennent des statistiques sur les messages d'erreur 404. Vérifier le nombre de liens rompus par ministère pourrait être un exercice utile, mais ces statistiques ne nous aideraient pas à fixer le niveau acceptable d'URL brisés. Faut-il ne rien tolérer ? Un pourcentage de 2 ou 3 %, par exemple, serait-il acceptable ? Pour répondre à cette question, il faudrait peut-être comprendre d'abord le motif expliquant la recherche de la permanence : l'objectif n'est pas de maintenir un URL pour ce qu'il est, mais plutôt de s'assurer que le document électronique situé à cette adresse pourra être trouvé et consulté, tant que le fichier électronique se trouve sur Internet. La question de savoir si les dossiers électroniques devraient être permanents (sur le Web ou ailleurs, sous leur forme originale ou autre) n'est pas étrangère à la gestion de l'information, mais dépasse le cadre du présent rapport.

Haut le page

B. Coûts et risques

- La vérification des liens par souci de permanence implique des coûts, engagés régulièrement par tous les ministères et agences du gouvernement. On peut aussi associer des coûts aux liens rompus. Ces coûts sont plus difficiles à quantifier, car il n'existe pas de mécanisme nous permettant de calculer le temps que consacrent les employés des ministères à chercher des documents. Cependant, même en supposant que l'ensemble du gouvernement fédéral et quelques ministères en particulier fassent preuve d'une plus grande efficacité à maintenir les URL que la communauté Internet en général, ces coûts ne sont pas négligeables. Bien sûr, qu'une personne passe quelques minutes à chercher un document, cela peut être négligeable. Cependant, si on les transpose à l'ensemble de la bureaucratie du gouvernement fédéral, ces quelques minutes deviennent rapidement des heures, d'autant plus qu'un nombre sans cesse croissant d'employés fédéraux sont vus comme des " travailleurs du savoir ", qui assument des fonctions entièrement axées sur l'information. Il importe aussi de garder en mémoire que l'information gouvernementale se fait toujours plus présente chaque jour sur le Web et que les passerelles augmentent en taille et en complexité.

- On ne peut calculer les coûts liés à la recherche d'information par les utilisateurs (publics), qui désirent savoir ce que fait leur gouvernement. Le public veut consulter l'information gouvernementale pour améliorer son existence et connaître les situations ou les événements susceptibles de poser un danger - mauvais temps, pollution atmosphérique, pneus défectueux -, pour ne nommer que ceux-là. Les gens d'affaires,

les universitaires et les chercheurs ont besoin de consulter rapidement des renseignements pour prendre des décisions éclairées. Il est impossible d'associer une valeur monétaire au temps qu'ils consacrent à chercher de l'information ou aux pertes (éventuelles) liées au fait de ne pas pouvoir la trouver. Même dans les cas où cela ne cause qu'un inconvénient ou un embarras mineur, il en résulte une érosion de la confiance dans le gouvernement avec le temps.

- Les pertes directes et la diminution de la productivité ne sont pas les seuls problèmes liés au manque de permanence des URL. Ce phénomène pose des risques plus graves, c'est-à-dire qu'il met en cause la capacité du gouvernement de prendre des décisions efficaces, lorsque celui-ci ne peut trouver ni utiliser efficacement sa propre information, et sa capacité à assumer comme il se doit sa responsabilité de rendre des comptes. La gestion des URL est un petit élément de la gestion de l'information, mais ce sont les différentes parties de ce tout qui déterminent la capacité du gouvernement à divulguer l'information qu'il a créée et à en permettre l'accès, et qui constitue le fondement même du mode de vie canadien. L'expérience individuelle de ne pouvoir consulter une ressource soit par ignorance, par erreur ou intentionnellement, a un effet négatif et, cumulativement, érode tout le système.

- Pour remplir leur rôle de gestionnaires de l'information, les employés fédéraux doivent faire preuve de leadership, disposer d'outils qu'ils comprennent et savent utiliser et posséder un formation adéquate.

Haut le page

C. Obstacles à une meilleure gestion des URL

- Il est certainement possible d'améliorer la gestion des URL en vue de promouvoir la permanence des ressources (du moins, dans certaines parties du gouvernement fédéral). Cependant, les URL ne fourniront jamais un identificateur unique indépendant de l'emplacement, de sorte qu'ils poseront toujours des difficultés en matière d'accessibilité.

- Malgré que divers systèmes généraux à identificateur unique aient vu le jour, il est évident que les systèmes de nom de domaine (URL) ont préséance en tant que principaux véhicules de localisation et de consultation de sources Web à court et à moyen termes.

- Il serait utile que les URL du gouvernement fédéral canadien soient " neutralisés ", de sorte qu'il ne soit plus nécessaire de les changer souvent, mais cette voie est impraticable, pour diverses raisons. L'élaboration et l'application de noms et de conventions de dénomination uniformes au niveau du sous-domaine auraient l'avantage de favoriser la cohérence. Parallèlement, on reconnaît la nécessité d'accommoder et de refléter les besoins de toutes les parties : ceux des utilisateurs (simplicité), ceux des développeurs (facilité d'encodage et de maintenance) et ceux des ministères et du gouvernement (possibilités de mise en valeur et promotion de l'image de marque).

Haut le page

D. La permanence n'est pas une priorité absolue

- Individuellement, les employés du gouvernement fédéral reconnaissent les problèmes causés par le manque d'URL permanents, mais le concept de permanence ne semble pas constituer une priorité pour l'ensemble des ministères et du gouvernement. La recherche et la mise en œuvre d'une solution passent d'abord par la sensibilisation au

problème et la reconnaissance qu'il s'agit d'une priorité. Dans le cadre décentralisé que nous connaissons présentement, faire de cette question une priorité va demander du temps, à moins que la Bibliothèque nationale du Canada et le Programme des services de dépôt n'acceptent d'assumer un rôle de leadership. L'importance grandissante d'une gestion de l'information efficace (y compris la gestion des URL) nécessite une collaboration et une coordination beaucoup plus étroites entre les organismes centraux qui assument une responsabilité directe en matière de gestion de l'information : le Conseil du Trésor, les Archives nationales, le greffier du Conseil privé, le Vérificateur général, la Bibliothèque nationale et le ministère de la Justice, pour ne nommer que ceux-là.

Haut le page

E. Le choix du moment pour implanter un système géré de localisateurs permanents

- Délaisser l'utilisation actuelle des URL au profit d'un système entièrement géré de localisateurs permanents impliquerait des coûts de lancement considérables en temps et en ressources et nécessiterait des investissements réguliers pour la maintenance annuelle. Pour le moment, l'adoption de l'un ou l'autre des systèmes gérés d'identificateurs permanents présentement disponibles poserait un risque moyen ou élevé, du fait que ces systèmes n'ont pas fait leurs preuves. Nous ne savons pas encore si aucun de ces systèmes ne sera jamais déployé à grande échelle. En outre, il n'est pas certain que les fureteurs conventionnels seront reconfigurés de façon à assurer la résolution de l'un ou l'autre de ces systèmes, sans l'ajout de modules d'extension.

- Le moment d'agir est un autre facteur à considérer. Il serait prématuré de tenter de passer à un système entièrement géré de localisateurs permanent sans effectuer le travail de fond qui s'impose. Dans la conjoncture actuelle, il est peu probable que cette approche bénéficie d'un soutien général au sein du gouvernement fédéral. Sans consensus, cette direction offre bien peu de perspectives de résultats à court et à moyen termes et constituerait un gaspillage de ressources déjà limitées. Cependant, les systèmes de localisateurs permanents existants méritent qu'on les surveille de près afin de déterminer si l'un ou l'autre sera largement utilisé sur le Web et de suivre l'évolution des mécanismes de résolution, surtout en ce qui concerne les capacités de résolution des URL offertes par les fureteurs conventionnels. Le fait de tenir l'ensemble du gouvernement fédéral davantage au courant de cette évolution mettra en lumière les lacunes des URL et contribuera à promouvoir la permanence.

- Le gouvernement dispose d'un cadre général permettant de promouvoir la permanence des ressources et d'en garantir l'accès. Le localisateurs permanents ne sont qu'un aspect de l'accès. Toute décision concernant la mise en œuvre d'URL permanents devrait être prise dans le contexte de la gestion du cycle de vie des ressources Web, des normes régissant la gestion du contenu Web et de la mise en œuvre de GED. C'est pourquoi il importe aussi que les décisions soient prises en collaboration avec d'autres intervenants, comme le Conseil du Trésor. La Bibliothèque nationale du Canada et le Programme des services de dépôt doivent s'efforcer d'établir des liens de collaboration avec d'autres parties à ce chapitre et participer aux projets existants axés sur des solutions pangouvernementales ou élaborer un projet applicable à l'ensemble du gouvernement.

- Tant que les URL ne seront pas remplacés par les URN ou les DOI, un autre identificateur général unique ou une autre technologie, il demeure possible de promouvoir la permanence et l'accessibilité de diverses manières : sensibilisation à l'importance de la permanence et à son rôle dans la mise en œuvre de GED, meilleure gestion des URL, promotion de l'archivage des dossiers électroniques comme élément crucial de la gestion du cycle de vie des ressources électroniques et collaboration active avec d'autres personnes et ministères investis de responsabilités en matière de gestion de l'information.

Haut le page

F. Besoin immédiat de leadership et de soutien financier

- Bien que les liens rompus et les fichiers électroniques inaccessibles ne soient pas les questions les plus populaires auprès des gestionnaires, la BNC et le PSD possèdent une expérience approfondie de la gestion de l'information qui les place en bonne position pour faire preuve de leadership en cette matière. À cette fin, la Bibliothèque nationale du Canada et le Programme des services de dépôt doivent, dans le contexte de GED, chercher à obtenir un soutien financier qui leur permettra de se doter des ressources nécessaires pour entreprendre des projets qui feront progresser la question et pour intervenir, en tant que participants ou promoteurs, dans des projets pilotes favorisant la permanence des publications du gouvernement fédéral sous forme électronique. Il serait possible de prendre part à des projets en voie d'élaboration par le Conseil du Trésor. La BNC et le PSD devraient aussi, dans une démarche prospective, s'informer d'autres possibilités de collaboration avant de lancer leur(s) propre(s) projet (s).

Haut le page

Recommandations

- La Bibliothèque nationale du Canada et le Programme des services de dépôt doivent examiner la question de la permanence et créer un groupe de travail ou une commission pour en faire un agent de promotion de la permanence. Le groupe de travail doit obtenir le droit de siéger au Comité consultatif du Secrétariat du Conseil du Trésor sur Internet afin d'assurer sa présence auprès des principaux décideurs et d'établir une voie lui permettant de communiquer ses connaissances spécialisées sur la question.

- Le groupe de travail devrait réunir des gestionnaires et des membres du personnel de différents ministères et organismes du gouvernement fédéral, œuvrant dans des domaines variés et aptes à présenter une perspective différente.

- Le groupe de travail devrait s'efforcer d'exercer une influence et de mettre à profit ses connaissances dans le but de promouvoir la permanence dans la gestion des ressources d'information électroniques du gouvernement fédéral. À cette fin, il doit :

Sensibilisation et promotion

- Favoriser la connaissance et la compréhension du problème, en le présentant comme une question à la fois technologique et sociale.
- Promouvoir la participation de la BNC et du PSD en tant que partenaires des ministères et autres organismes fédéraux pour la promotion de la permanence et de l'accessibilité des ressources électroniques.
- Promouvoir une meilleure gestion des URL et le partage de l'information sur les " pratiques exemplaires ".
- Promouvoir l'adoption d'une déclaration sur la permanence.

Renforcement des normes et des lignes directrices fédérales

- Améliorer l'utilité des normes et lignes directrices fédérales existantes afin qu'elles soient mises en œuvre de façon uniforme au gouvernement et que les employés possèdent les connaissances et les outils leur permettant de bien gérer l'information électronique

- Renforcer les normes et lignes directrices fédérales pertinentes en comblant leurs lacunes et en y apportant des révisions et des mises à jour au besoin. Par exemple, élaborer une norme de communications et modifier les lignes directrices concernant les " vanity URL " [adresses URL personnalisées] ; envisager d'établir des lignes directrices pour les noms de fichiers des sous-domaines et des sous-répertoires.

Recherche de financement pour la participation à des projets pilotes

- Examiner les possibilités de participer à des projets pilotes qui visent à promouvoir la permanence, qui sont applicables à l'ensemble du gouvernement et qui améliorent la prestation des services du gouvernement fédéral en direct. Si un projet pertinent est proposé,[40] : demander une aide financière en vertu du programme GED. Sinon, élaborer une analyse de cas afin d'obtenir le financement d'un projet mise en œuvre conjointement par la BNC et le PSD. Il est recommandé que la BNC et le PSD envisagent de mener un projet évaluant le concept d'un répertoire national de ressources d'information choisies utilisant un modèle physiquement décentralisé et un système à code d'identification unique.

Haut le page

Notes de conclusion

1 Pour obtenir un complement d'information, voir l'article « A Metadate Kernel for Electronic Permanence », dans la bibliographie.

2 Pour obtenir de plus amples détails, voir la norme provisoire intitulé Dynamic Delegation Discovery System (DDDS), de M. Mealing, dans la bibliographie.

3 Organization for the Advancement of Structured Information Standards, www.oasis-open.org.

4 Voir les observations à ce sujet dans un rapport australien intitulé « Persistent Identification Systems », Partie 1, section 1.3.4

5 Pour en savoir davantage, voir les articles de Juha Hakala et l'enregistrement des espaces de noms au site Web de l'Université d'Helsinki mentionné dans la bibliographie.

6 Voir le rapport du GPO intitulé « Persistent Identification Systems », dans la bibliographie.

7 GAO, Document 01-428, « Information Management: Electronic Dissemination of Government Publications », Volume 1, mars 2001.

8 Voir le site Web www.purl.oclc.org/problems.html.

9 Voir quelques éléments de documentation à ce sujet dans la bibliographie, à la section consacrée au système Handles. Pour en savoir davantage sur la dénomination, voir le document intitulé « LC Naming Conventions for Digital Resources », sur le site Web de la Library of Congress Digital Library.

10 Voir, dans la bibliographie, l'article intitulé « A Test Bed for Digital Object

Registration and Management: The Copyright Office Electronic Registration, Recordation and Deposit System: 2000 Project Summary ».

11 Voir l'article intitulé « Some Applications of the Handle System », dans la bibliographie.

12 Il s'agit d'une entreprise de sécurité privée, selon une information fournie par courrier électronique par Jane Euler, le 30 janvier 2002.

13 Par exemple, l'Australie a mené un essai pilote sur le Système Handles (voir l'annexe 2 du rapport intitulé « Persistent Identification Systems ». Les autorités australiennes ont établi que bien que le système présente la fonctionnalité technique fondamentale pour une bibliothèque numérique et qu'il représente la meilleure option présentement disponible, elles ne pouvaient le recommander en raison de l'investissement requis pour la mise en œuvre du système et des marches à suivre ainsi que des modifications qu'il fallait apporter aux systèmes. En outre, le système Handles nécessite une connaissance approfondie d'UNIX et de Java.

14 IDF a négocié des niveaux de soutien et des normes de service garantis avec la CNRI, qui a confié le soutien du système à une entreprise commerciale et garantit maintenant un fonctionnement continu.

15 Voir l'article intitulé « Overview of Handles System », dans la bibliographie.

16 CrossRef sert le secteur de la publication de revues ; Content Direction et Enpia Systems (Corée) servent le secteur du livre électronique et Learning Object Network prévoit enregistrer les DOI dans le contexte de la gestion des éléments d'apprentissage.

17 Le modèle commercial demeure incertain. Toutefois, dans le modèle original, il en coûte 1 000 $ US pour un préfixe permettant la production d'un nombre indéterminé de DOI. On prévoit que le prix du préfixe variera selon l'autorité d'enregistrement et la diversité des services qu'elle offre.

18 Voir les exemples de DOI utilisés par le système IDEAL d'Academic Press, à www.apnet.com/.

19 Voir l'article intitulé « The DOD Goes DOI: Massive Project May Put Wheels Under DOI Initiative ».

20 Voir l'article intitulé « DOI: A Tale of Whoa », dans la bibliographie.

21 Voir, dans la bibliographie, l'article intitulé « Reference Linking With DOI: A Case Study », qui porte sur l'applicabilité du système à l'établissement de liens de référence.

22 Pour le moment, les Presses établissent des liens vers l'intérieur, mais pas de liens vers l'extérieur, mécanisme conçu pour proposer aux utilisateurs des liens vers d'autres ressources Web pertinentes

23 Par exemple, en collaboration avec d'autres organisations, CrossRef a démontré qu'il était possible d'établir des « liens localisés », qui dirigent les chercheurs vers une version d'un article offrant un accès au texte intégral sans qu'ils aient à payer (par exemple, un abonnement à une bibliothèque locale) plutôt que vers une version nécessitant un déboursé.

24 Microsoft a décidé d'utiliser les DOI dans son lecteur eBook. On peut consulter quelques scénarios de démonstration www.seyboldreports.com/ebooks/features/011010-doi.html.

25 Selon un courrier électronique de J. Kunze, développeur de ARK. (Courriel en date du 2 octobre 2002).

26 Selon un courriel signé de J. Kunze, développeur du système ARK, pour le moment, la mise en œuvre du système se limite à la production d'environ six pages de codes PHP utilisant l'Entrez API à la base de données Pubmed, illustrant le protocole et la structure de métadonnées décrits dans les données techniques du système ARK. « Le difficile travail d'assigner des codes de permanence n'a pas été fait. » (Courriel daté du 24 janvier 2002).

27 Voir, dans la bibliographie, l'article de John Kunze intitulé « A Metadata Kernel for Electronic Permanence ».

28 La bibliographie énumère toute l'information disponible sur cet outil.

29 D'après un courriel reçu de Victoria Reich, gestionnaire de projet, en date du 27 janvier 2002.

30 Pour en savoir davantage, visitez le site Web www.openly.com/.

31 Également utilisé par les Presses scientifiques du CNCR.

32 Utilisent une technologie OpenURL, service de résolution qui prend en compte l'information locale de l'utilisateur au moment de le diriger vers une ressource. Pour en savoir davantage, se référer aux articles de la bibliographie. DOI considère que la technologie OpenURL est complémentaire. Selon la NISO, pour le moment, on ne peut définir clairement la relation à établir entre la norme OpenURL et les normes des métadonnées et celles des codes d'identification.

33 Voir, dans la bibliographie, l'article d'Andy Powell intitulé « Encoding OpenURLs in Dublin Core Metadata ».

34 Communiqué de presse de la NISO daté du 11 février 2002. Pour en savoir davantage, voir http://library.caltech.edu/openurl/default.htm .

35 Voir les normes de W3C pour RDF / XML. Voir aussi les normes de W3C pour Xpath, logiciel d'adressage de documents XML.

36 Peu de sources documentaires donnent des indications précises sur l'application concrète de cette approche. Il semble que des éditeurs scientifiques allemands aient effectué des travaux dans ce domaine, mais on ne dispose pas de détails pour le moment, car tous les documents sources découverts sont rédigés en allemand (voir le projet CARMEN à www.bis.uni-oldenburg.de/carmen_ap4/). Le milieu des bibliothèques du R.-U. a aussi examiné le RFD dans le contexte qui lui est propre. Voir les notes du UKOLN RDF Seminar au site Web www.ukoln.ac.uk.

37 D'après un message électronique de M. Byrnes, président du Groupe de travail sur la permanence de l'information électronique, établi par la Bibliothèque (daté du 25 janvier 2002).

38 On peut prendre connaissance de cette version condensée dans l'article intitulé « Defining NLMs Commitment to the Permanence of Electronic Information », p. 3, cité dans la bibliographie.

39 Voir le document intitulé « Safeguarding Australia's Web Resources: Guidelines for Creators and Publishers », cité dans la bibliographie.

40 Des renseignements provisoires indiquent qu'il y aurait une possibilité de participer

à un projet sur les métadonnées amorcé par le Conseil du Trésor, mais les détails restent inconnus pour le moment. Le Groupe de travail devrait saisir cette occasion, dans un esprit prospectif. Si ce projet ne convient pas, le Groupe devrait rechercher d'autres possibilitsé avant d'amorcer un projet de sa propre initiative.

Haut le page

Création : 1996-05-22
Mise à jour : 2004-01-12

Haut de la page **Avis importants**

Localisateurs permanents pour les publications du gouvernement fédéral canadien

Bibliographie

N.B. Cette bibliographie, compilée en février et mars 2002, est partiellement annotée. Certaines liens ne fonctionne plus.

- Générale
- Identificateurs permanents - systèmes
  - Le système Handle
  - Système DOI
  - Système ARKs
  - PURLs
- URLs, URNs et URIs
- URNs et NBNs
- Autres espaces de noms
- Systèmes d'indentification des Éditeurs
- Permanence dans les citations de revues
- Ressources Australiennes sur la question de la permanence
- Organisations
- OpenURL
- Divers

Rapport - Table des matières

Générale

Arms, William. « Key Concepts in the Architecture of the Digital Library », D-Lib Magazine, juillet 1995. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/July95/07arms.html

Caplan, Priscilla et Flecker, D. Choosing the Appropriate Copy, Digital Library Federation Architecture Committee. septembre 1999. Réference du 5 mars 2002.
URL : http://www.niso.org/DLFarch

Causton, Laurie. Identifying and Describing Web Resources, (Commission européenne DGXIII/E-4). Réference du 5 mars 2002.
URL : http://www.elpub.org/html/webres.html

CrossRef : The Central Source for Reference Linking. Réference du 5 mars 2002.
URL : www.crossref.org

Doyle, Mark. « Citing and Linking in Electronic Scholarly Publishing: a Pragmatic Approach ». American Physical Society. [1999?]. Réference du 5 mars 2002.
URL : http://ridge.aps.org/DOYLE/sweden/talk/sweden.html

Engelbart, Douglas C. « Knowledge-Domain Interoperability and an Open Hyperdocument System », juin 1990. (AUGMENT, 132082,). Réference du 5 mars 2002.
URL : www.bootstrap.org/augment-132082.htm

Flecker, Dale. Citation - E-Journal Links from the (Academic) Library Viewpoint: A presentation to the NISO/NFAIS/SSP Linking Workshop, 11 février 1999, Washington, DC. Réference du 5 mars 2002.
URL : www.niso.org/presentations/flecker.html

Green, Brian et Bide, Mark. Unique Identifiers: A Brief Introduction, (sans date, probablement en 1997). Réference du 5 mars 2002.
URL : www.bic.org.uk/uniquid.html

Guenther, Rebecca. Naming Conventions for Digital Resources. Library of Congress Network Development and MARC Standards Office. Réference du 5 mars 2002.
URL : http://lcweb.loc.gov/marc/naming.html

Information Object Numbering Systems (IONS) : JTAP report 42. Division of Learning Development, De Monfort University. Réference du 5 mars 2002.
URL : http://www.jtap.ac.uk/reports/htm/jtap-042.html#_Toc454868885

Kahn, Robert et Robert Wilensky. « Framework for Distributed Digital Object Services ». 13 mai 1995. cnri.dlib/tn95-01. Réference du 5 mars 2002.
URL : www.cnri.reston.va.us/k-w.html

Koehler, Wallace. « Digital Libraries and World Wide Web Sites and Page Persistence ». University of Oklahama, 2000. Réference du 5 mars 2002.
URL : http://informationr.net/ir/4-4/paper60.html
Cet article examine les notions de permanence des pages et leur intermittence, décrite comme la disparition puis la réapparition de documents Web. À tout moment, environ 5 % des pages Web sont intermittentes, ce qui signifie qu'elles ne sont pas disponibles, mais qu'elles le redeviendront. D'après une étude longitudinale de 120 semaines effectuée sur un échantillon de documents Web, il semble que la période de demi-vie d'une page Web soit de moins de deux ans et que celle d'un site Web soit de plus de deux ans. Cela veut dire qu'une collection de documents Web non expurgée, créée il y a deux ans, contiendrait le même nombre d'URL, mais que seulement la moitié des URL mèneraient à un contenu.

Miller, James S. « W3C and Digital Libraries », D-Lib Magazine, novembre 1996. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/november96/11miller.html

Naming and Linking Strategy. RLG Digital Collections Project. Studies in Scarlet Project. Réference du 5 mars 2002.
URL : http://www.rlg.org/scarlet/name.html

Nelson, Michael L. et B. Danette Allene. « Object Persistence and Availability in Digital Libraries ». D-Lib Magazine, janvier 2002. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/january02/nelson/
01nelson.html
Cet article analyse un sondage sur la permanence et la disponibilité de 1 000 objets de bibliothèque numérique. À partir de 151 échantillons de données, l'étude conclut que 3 % de tous les objets n'étaient plus disponibles, ce qui correspond au nombre d'URL perdus établis par Lawrence et al. dans leur étude de 2001 intitulée.

Nelson, Michael L. et al. « Persistence of Web References in Scientific Research ». IEEE Computer, 34, (2), 26-31. Réference du 5 mars 2002.
URL : http://csdl.computer.org/comp/mags/co/2001/02/r2026abs.htm

The Open Archives Initiative Protocol for Metadata Harvesting. 2.3. Unique Identifiers - Document Version 2001-04-24. Réference du 5 mars 2002.
URL : www.openarchives.org/OAI/
openarchivesprotocol.htm#Unique Identifier

Paskin, Norman. « E-Citations : Actionable Identifiers and Scholarly Referencing ». Version 1.2, 17 décembre 1999. Réference du 5 mars 2002.
URL : http://dx.doi.org/10.1000/170

Paskin, Norman. Information Identifiers. Learned Publishing. Vol 10, no. 2, 135-156. avril 1997. Réference du 5 mars 2002.
URL : http://www.elsevier.com/inca/homepage/
about/infoident/Menu.shtml"

Paskin, Norman. Position paper for W3C Workshop on Digital Rights Management for the Web (22/23 janvier 2001) Réference du 5 mars 2002.
URL : www.doi.org/001219_IDF_DRM_pos_paper.htm

Payette, Sandra. « Persistent Identifiers on the Digital Terrain ». RLG DigiNews. 15 avril 1998. Réference du 5 mars 2002.
URL : www.rlg.org/preserv/diginews/
diginews22.html#Identifiers
Cet article décrit clairement les URN et les principaux exemples de mise en œuvre d'identificateurs permanents, dont les PURL, le Handles System et les DOI. Il examine aussi brièvement le processus de mise en œuvre des identificateurs permanents pour les projets d'imagerie numérique.

Requirements for an Identification Scheme for BIBLINK. mai 1977. BIBLINK - LB 4034, Section D4.1 Identification. Réference du 5 mars 2002.
URL : http://hosted.ukoln.ac.uk/biblink/wp2/d2.1/doc0004.htm

Sollins, Karen R. et Jeffrey R. Van Dyke. Linking in a Global Information Architecture. Référence du 5 mars 2002
URL : http://ana-www.lcs.mit.edu/anaweb/links.html

Universal Document Naming Systems. Réference du 5 mars 2002.
URL : http://elib.cs.sfu.ca/USIN/naming.html

« Use of DNS Aliases for Network Services », RFC 2219, 1997. Réference du 5 mars 2002.
URL : www.dns.net/dnsrd/rfc/rfc2219.html
Pratiques exemplaires Internet actuelles

Web Pages Must Live Forever. Jakob Nielsen's Alert box. 29 novembre 1998. Réference du 5 mars 2002.
URL : www.useit.com/alertbox/981129.html

Haut de la page

Identificateurs permanents - systèmes

Le système Handle

A global system of Handle Servers. Library of Congress, NDLP (National Digital Library Program) Internal documentation. Réference du 5 mars 2002.
URL : http://lcweb2.loc.gov/ammem/award/docs/h-s3.html

Grass, Judith et William Y. Arms. A Syntax for Distributed Object Handles. 3 août 1994. Réference du 5 mars 2002.
URL : www.ifla.org/documents/libraries/
cataloging/metadata/hdl_form.txt
Ce document aborde certaines questions soulevées par la conception d'une structure de pointeurs et en expose la syntaxe. Cette description de la syntaxe des pointeurs est conforme aux exigences de la DLS et permet une utilisation facile des pointeurs d'objets avec un grand nombre d'outils de navigation Internet existants.

« Handle Server Overview ». Library of Congress National Digital Library Program Internal Documentation. 1997. Réference du 5 mars 2002.
URL: http://lcweb2.loc.gov/ammem/award/docs/
handle-server.html
Aperçu du système Handle

Handle System Namespace and Service Definition. IETF Internet Draft. avril 2001. Réference du 5 mars 2002.
URL : www.idrm.org/draft-irtf-idrm-handle-system-def-01.txt
Description détaillée de l'espace de noms, des données et du modèle de service Handles, ainsi que de son protocole d'opération et d'authentification.

More on handles and naming authorities. Library of Congress, NDLP (National Digital Library Program) Internal documentation. Réference du 5 mars 2002.
URL : http://lcweb2.loc.gov/ammem/award/docs/h-s2.html

Sam X. Sun. « Establishing Persistent Identity Using the Handle System ». Corporation for National Research Initiatives. 2001. Réference du 5 mars 2002.
URL : www10.org/cdrom/posters/1136.pdf
Cet article établit que la référence d'identité utilisée pour le système de communication réparti devrait être définie sans référence aux caractéristiques ni aux clés publiques associées au sujet sous-jacent, permettant ainsi d'assurer la permanence de la référence d'identité par rapport aux changements apportés aux caractéristiques ou aux clés publiques. L'auteur propose de séparer l'établissement de la confiance à l'égard de la distribution répartie en trois catégories : fiabilité du transport, fiabilité de l'administration et crédibilité du contenu, ce qui simplifie la question de la confiance et sa gestion. Cela rejoint notre proposition d'utiliser les pointeurs inscrits au système Handles en tant que références d'identité permanentes pour la communication Internet, et d'utiliser le service de sécurité du système de pointeurs pour appuyer la gestion de la confiance sur Internet.

Sun, Sam X et al. « Handle System Protocol (ver 2.0) Specification ». IETF Internet Draft. avril 2001. Réference du 5 mars 2002.
URL : www.idrm.org/
draft-irtf-idrm-handle-system-protocol-01.txt
Description du protocole du Système Handles, version 2, pour Internet.

Sun, Sam X et al. Overview of Handles System. août 2000. Réference du 5 mars 2002.
URL : http://www.handle.net/overview-current.html

« Technical Information on Digital Objects, Handles and Repositories ». Corporation for National Research Initiatives, 1997. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/february97/cnri/
02arms2.html

Haut de la page

Système DOI

Arms, William Y. « Digital Object Identifiers (DOIs) and Clifford Lynch's Five Questions on Identifiers ». 13 octobre 1997. Réference du 5 mars 2002.
URL : www.arl.org/newsltr/194/arms.html
Cet article livre une information documentaire sur les DOI dans le contexte des cinq questions que pose Clifford Lynch sur les systèmes d'identificateur dans son article intitulé «Identifiers and Their Role In Networked Information Applications ».

Atkins, Helen et al. « Reference Linking with DOIs : a Case Study ». D-Lib Magazine, février 2000. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/february00/02risher.html

Bide, Mark. « In Search of the Unicorn : The Digital Object Identifier from a User Perspective ». février 1998. Réference du 5 mars 2002.
URL : www.bic.org.uk/unicorn2.pdf
Cet article de 40 pages explique pourquoi nous avons besoin de nouveaux identificateurs uniques et examine les premiers efforts d'élaboration des DOI. L'auteur analyse les besoins particuliers de différents utilisateurs. Information documentaire utile pour comprendre la genèse des DOI et la complexité des différents besoins des utilisateurs.

« The DOD Does DOI : Massive Project May Put Wheels under DOI Initiative ». 16 octobre 2001. Réference du 5 mars 2002.
URL : www.seyboldreports.com/ebooks/news/011016-dod.html
Communiqué sur la décision du Pentagone de convertir tous ses dossiers papier en formats électroniques standard et de les cataloguer au moyen de DOI. Bien que les trois organismes d'enregistrement actuels n'aient enregistré que quelques millions de DOI, ce nombre devrait se chiffrer dans les billions avec le projet du DOD.

The DOI Handbook. Version 1.0.0, février 2001. Réference du 5 mars 2002.
NB: Version 3.2, août 2003 disponible à
URL : a href="http://www.doi.org/handbook_2000/index.html
Ce manuel livre une information détaillée sur le système DOI, son histoire et son élaboration, ainsi que sur ses composantes (énumération, description, résolution, élaboration de politiques). Il comprend aussi des chapitres sur le rôle de l'organisme de maintenance, la nomination et le rôle des organismes d'enregistrement et la gouvernance. Les annexes livrent de l'information sur la syntaxe des DOI, le système Handles et le cadre de métadonnées ainsi qu'un aperçu des profils d'application. Le document comprend des renseignements pratiques sur la façon d'obtenir un préfixe DOI, de choisir un organisme d'enregistrement et de devenir membre de la DOI Foundation.

DOI: The Digital Object Identifier System. Developed by the International DOI Foundation. Réference du 5 mars 2002.
URL : www.doi.org/overview/sys_overview_021601.html
Aperçu à jour des DOI

Felici, Jim. « DOI: A Tale of Whoa ». The Seybold Report, 19 novembre 2001. Réference du 5 mars 2002.
URL : www.seyboldreports.com/TSR/subs/0116/
html/news-doi.html
NB : Nom d'utilisateur et mot de passe sont requise
L'auteur examine brièvement les DOI et conclut que bien qu'on n'entrevoie encore aucune technologie qui soit comparable aux DOI, ce système n'a pas fait l'unanimité

chez les éditeurs. Le soutien à l'élaboration de normes par les droits d'adhésion est plutôt faible. Les secteurs de la musique et du film ont élaboré leur propre système ; pour leur part W3C et IETF n'y ont pas montré beaucoup d'intérêt. Un des problèmes est qu'il ne s'agit pas d'un groupe de normes, à proprement parler. Par ailleurs, le système des DOI est mal compris et IDF n'a pas les moyens de mieux le faire connaître. Giant Chair, intervenant de premier ordre dans le secteur du livre électronique, a rejeté le système après l'avoir examiné. McGraw-Hill a fait une publicité encourageante au sujet de DOI dans ses offres et le secteur des revues techniques et scientifiques a adopté ce concept. Cependant, peu d'autres éditeurs ont emboîté le pas.

« How the "Digital Object Identifier" Works ». Business Week. 23 juillet 2001. Réference du 5 mars 2002.
URL : www.businessweek.com/magazine/content/01_30/b3742032.htm
Articles générales sur les DOI.

IDF Terms of Governance and Policies with Registration Agencies, Version 4, 18 avril 2001. Réference du 5 mars 2002.
NB: Version 5, janvier 2002 disponible à
URL : www.doi.org/doi-ap/terms.pdf
Ce document de 17 pages expose en détail le mode de gestion et la politique d'IDF au regard des organismes d'enregistrement et établit la situation relative à la politique. On formule des recommandation. dans le cas où la politique n'a pas fait consensus.

International DOI Foundation. Annual Review, septembre 2000. Réference du 5 mars 2002.
URL : www.doi.org/idf-ann-rpt-10-00.pdf
Ce rapport de 36 pages expose la situation et l'évolution des DOI en 2002. Comprend une liste des alliances stratégiques et des membres par catégorie. Comprend le rapport Treasurers.

Paskin, Norman. DOI : An Overview of Current Status and Outlook. avril 1999, The International DOI Foundation. Réference du 5 mars 2002.
URL : http://www.icsti.org/forum/fo9904.html/#paskin
Aperçu du système DOI en date d'avril 1999

Paskin, Norman. « Digital Object Identifier: Implementing a standard digital identifier as the key to effective digital rights management ». DOI, avril 2000. Réference du 5 mars 2002.
URL : www.doi.org/doi_presentations/aprilpaper.pdf
Fondements du système DOI et analyse de cas.

Paskin, Norman. « Digital Object Identifier: The Case for Supporting the International DOI Foundation », 2001. Réference du 5 mars 2002.
URL : www.doi.org/doi_presentations/010808.pdf
Présentation de 20 pages sur les caractéristiques et les avantages du système DOI.

Paskin, Norman. « DOI Commercial Implementations: DOI Update 2001 », Diaporama, 2001. Réference du 5 mars 2002.
URL : www.doi.org/frank_bf_oct01/011010Frankfurt.htm
Présentation de 62 dispositives sur les DOI, dans le contexte de mises en œuvre commerciales - CrossRef et Learning Systems.

Paskin, Norman et al. « DOI Implementations », Diaporama, 2001. Réference du 5 mars 2002.
URL : www.doi.org/frank_bf_oct01/011010Frankfurt2.htm
Présentation de 14 dispositives sur les DOI, l'édition de livre et les objets d'apprentissage.

Paskin, Norman. « E-Citations: Action Identifiers and Scholarly Publishing ». juillet 2000. Réference du 5 mars 2002.
URL : www.doi.org/citations4.pdf
Cet article examine le rôle des « identificateurs aptes à exécuter une action » [actionable identifiers], comme les DOI, au regard des citations savantes dans un environnement numérique. Le consortium CrossRef est un exemple concret d'utilisation des DOI faisant présentement l'objet de travaux.

Paskin, Norman et al. « From One to Many: The Next Stage in the Development of DOI Functionality ». A DOI Discussion Paper. août 2000. Réference du 5 mars 2002.
URL :
URL : www.doi.org/one2many.pdf
Décrit les plans et les activités permettant d'étendre considérablement la fonctionnalité des DOI. Il s'agit d'un des aspects en voie de développement, les autres étant la mise en œuvre des structures de métadonnées et le déploiement commercial des DOI par le biais d'organismes d'enregistrement. Un DOI est un identificateur apte à exécuter une action : le fait de cliquer sur un DOI commande une action. Dans la mise en œuvre initiale, il s'agit d'une action simple : un clic sur un DOI mène l'utilisateur à un URL (« un à un »). Une action « un à plusieurs » suppose qu'un seul DOI peut servir à exécuter plus d'une action. C'est là l'objectif de la mise en œuvre intégrale des DOI. Ce document examine les mécanismes menant d'une action « un à un » à une action « un à plusieurs » et les conséquences de cette transformation.

Powell, Andy. « Resolving DOI based URNs Using Squid », D-Lib magazine, 15 juin 1998. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/june98/06powell.html
Cet article décrit un système expérimental permettant que des DOI chiffrés comme URN soient résolus à la place des fureteurs Web. Il livre aussi de l'information sur les questions liées aux identificateurs à long terme et les perspectives possibles dans ce domaine.

« Reference Linking with DOIs : A Case Study ». D-Lib Magazine. février 2000. Réference du 5 mars 2002.
URL : www.dlib.org.ar/dlib/february00/02risher.html"
Cet article aborde la création de DOI-X, prototype de base de métadonnées conçu pour soutenir les recherches de DOI et qui assure l'intégration de l'enregistrement et de la maintenance des métadonnées, permettant aux éditeurs d'utiliser un seul mécanisme et un seul processus d'assurance de la qualité pour enregistrer les DOI et les métadonnées qui y sont associées. Il contient aussi les mécanismes de recherche nécessaires pour accéder aux métadonnées des articles de journal, tant par recherche d'article unique que par recherche en lot, de façon à faciliter l'établissement de liens de référence. Le prototype de base de données a fait l'objet d'une démonstration en octobre 1999.

Sidman, David et Tom Davidson. A Practical Guide to Automating the Digital Supply Chain with the Digital Object Identifier (DOI). Content Directions, Inc., 2001. Réference du 5 mars 2002.
URL : www.doi.org/topics/CDI-PRQ2001-article.pdf
Cet article examine la similitude entre les facteurs agissant sur l'adoption du système DOI et celui de l'UPC. La principale valeur des deux systèmes est leur rôle en tant qu'identificateur de produit universel pour toute la chaîne d'approvisionnement ; cette valeur ne se manifeste que lorsqu'une masse critique d'intervenants adoptent cette norme. Le système DOI présente les mêmes perspectives de succès en offrant un rendement rapide et indépendant, sur le plan des recettes, en faisant en sorte qu'il soit plus facile pour les éditeurs d'être payés par la voie de distribution numérique et, sur le plan des coûts, en limitant l'investissement que doivent consentir les éditeurs pour entrer de plein pied dans le secteur de l'édition électronique.

Technical Note : DOI and OpenURL. DOI, 2001. Réference du 5 mars 2002.
URL : www.doi.org/topics/010402DOI_OpenURL.pdf
Document de 4 pages examinant les DOI et les OpenURL, qui sont des technologies

complémentaires : le système DOI crée un identificateur résolvable mondialement accessible qui peut être maintenu indépendamment des liens emboîtés ; OpenURL offre un service de résolution qui tient compte de l'information de l'utilisateur local au moment de le diriger vers une ressource.

Haut de la page

Système ARKs

Kunze, J. (2001) The ARK Persistent Identifier Scheme, IEFT working draft, 7 septembre 2001. Réference du 5 mars 2002.
URL : http://www.ietf.org/internet-drafts/draft-kunze-ark-02.txt

Kunze, John. (2001). A Metadata Kernel For Electronic Permanence. Journal of Digital Information. 2:2; 11 juin 2001. Réference du 5 mars 2002.
URL : http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Kunze
La permanence de l'information électronique est liée à l'identification permanente et à certains types de métadonnées. Différents systèmes de dénomination ont été proposés pour permettre l'identification permanente - Nom de ressource uniforme, Identificateur d'objet numérique et URL permanent -, chacun intégrant une forme de séquence de caractères reconnaissables à l'identificateur afin d'indiquer au récepteur l'intention de l'autorité d'attribution de nom de le faire perdurer indéfiniment. La Clé de ressource archivistique (Archival Resource Key - ARK) est un système de dénomination permettant des références aux objets qui sont durables à long terme, fondé sur le principe que la permanence est uniquement une question de service. La permanence n'est pas inhérente à un objet, pas plus qu'elle n'est un caractère conféré à l'objet par une syntaxe ou un préfixe de dénomination. Elle est assurée par la gestion efficace des objets et de leurs identificateurs par le fournisseur. ARK est un identificateur qui lie un objet ainsi que les métadonnées portant l'engagement d'un fournisseur donné et une description de l'objet afin d'assurer une identification complète. De nombreux protocoles et structures d'enregistrement peuvent assurer le service des ARK, mais la structure simple et générale appelée Electronic Resource Citation (ERC) (Référence de ressource électronique) répond aux exigences de services de permanence réalistes. À partir du consensus de Dublin Core, ERC établit un sous-ensemble de quatre cases sémantiques - noyau de métadonnées -, qui répond à la nécessité d'identifier adéquatement les objets permanent et de produire des métadonnées à faible coût.

Haut de la page

PURLs

Library of Congress. National Digital Library Program. (1997). The Relationship between URNs, Handles and PURLs. 12 août 1997. Réference du 5 mars 2002.
URL : http://lcweb2.loc.gov/ammem/award/docs/PURL-handle.html
Cet article décrit les URN, le système Handles et les PURL et donne des exemples sur leur fonctionnement. Il aborde brièvement l'interfonctionnement des pointeurs et des PURL, les résolveurs PURL agissant comme serveurs mandataires des pointeurs. Conclusions : (1) Pour une institution désireuse d'assurer la permanence des identificateurs pour les ressources qu'elle possède et gère, utiliser un résolveur PURL peut s'avérer uns solution pratique. (2) Quelle que soit la forme d'URN ou de PURL, on doit assumer la responsabilité de maintenir leurs URN. Cette tâche est beaucoup moins lourde que d'effectuer le suivi et les modifications de chaque lien en ligne. Les

systèmes Handles et PURL comprennent des outils de mise à jour des dossiers.

OCLC Online Computer Library Center, Inc. (2000). Known problems in the OCLC PURL Resolver software release 2.14. Dublin, OH: OCLC Online Computer Library Center, Inc. Réference du 5 mars 2002.
URL : www.purl.oclc.org/problems.html
Page sur les problèmes associés à Resolver Software 2.1.4.

OCLC Online Computer Library Center, Inc. (2002). PURL. Dublin, OH: OCLC Online Computer Library Center, Inc. Réference du 5 mars 2002.
URL : www.purl.oclc.org
Page d'accueil de l'OCLC sur les on PURL.

Shafer, Keith et al. (1996?) Introduction to PURLs. Dublin, OH : OCLC Online Computer Library Center, Inc. Réference du 5 mars 2002.
URL : http://purl.oclc.org/OCLC/PURL/INET96

Weibel, Stuart et al. (2002). PURLs: Persistent Uniform Resource Locators. OCLC Online Computer Library Center, Inc. Réference du 5 mars 2002.
URL : http://purl.oclc.org/OCLC/PURL/SUMMARY
Survol rapide des PURL.

Haut de la page

URLs, URNs et URIs

Berners-Lee, T. (1998). Uniform Resource Identifiers (URI) : generic syntax. RFC 2396. Internet Engineering Task Force (IETF). Network Working Group. Réference du 5 mars 2002.
URL : www.ietf.org/rfc/rfc2396.txt?number=2396

Hakala, J. (1999). Uniform Resource Names. Tietolinja News, 15 mars 1998. Accessed mars 5, 2002.
URL : www.lib.helsinki.fi/tietolinja/
0199/urnart.html
Cet aperçu des URN comprend une description sommaire de leur fonctionnement et de leur capacité de s'adapter à d'autres identificateurs ; il aborde aussi l'utilisation d'espaces de noms, y compris l'usage que peuvent en faire les organismes de l'ISBN, l'élaboration d'un service de résolution URN et la décision de la Conference of European National Librarians (CENL) de mettre en œuvre les URN.

Hakli, Esko et J. Hakala. (1998). URN implementation in National Libraries. Helsinki University Library, 1 octobre 1998. Réference du 5 mars 2002.
URL : http://linnea.helsinki.fi/urn/urnimp.html«
Aborde des questions concernant la mise en œuvre d'URN par des bibliothèques nationales et donne des recommandations à la Conference of European National Librarians (CENL). Le document explique pourquoi les bibliothèques nationales devraient mettre en œuvre les URN. Il examine les coûts possibles, établit une comparaison entre URN et DOI et donne la liste des documents de spécifications du groupe de travail de l'IEFT sur l'URN.

Jones, Dominic. (2001). Avoiding 'Linkrot' on your site. IR Web Report. 12 décembre 2001. Réference du 5 mars 2002.
URL : www.irwebreport.com/features/010703.htm

Library of Congress. National Digital Library Program. (1997). Examples of

mechanisms for keeping URLs relatively persistent. Appendix to : Arms, Caroline R. Access aids and interoperability. (1997). Washington, DC : Library of Congress. National Digital Library Program. Réference du 5 mars 2002.
URL : http://lcweb2.loc.gov/ammem/award/docs/persistence.html
Utilisation des pseudonymes, scripts cgi-bin, modes de dénomination logiques, utilisation de résolveurs URL indépendants ou d'un résolveur mandataire de pointeurs pour assurer la permanence. Exemples concrets.

Moats, R. (1997). URN syntax. RFC 2141. Internet Engineering Task Force (IETF). Network Working Group, mai 1997. Réference du 5 mars 2002.
URL : www.ietf.org/rfc/rfc2141.txt?number=2141

Peacock, Ian. (1998). What is a...URI? Ariadne. 18 : 1 décembre 1998. Réference du mars, 2002.
URL : www.ariadne.ac.uk/issue18/what-is
Initiation aux URI, aux URL, aux URN et aux URC.

Van der Werf-Davelaar, Titia. (1999). Identification, location and versioning of web-resources. URI Discussion paper. Réference du 5 mars 2002.
URL : www.konbib.nl/donor/rapporten/URI.html

Walsh, Norman. (2001). A URN namespace for public identifiers. Cover Pages, 4 mai 2001. (draft-urn-publicid-03) . Réference du 5 mars 2002.
URL : http://xml.coverpages.org/walshPINamespaceAnn.html
Cet article décrit un espace de noms URN conçu pour permettre que des identificateurs publics soient exprimés dans la syntaxe URI. Les entités externes XML ont deux identificateurs : un identificateur système et un identificateur public. L'identificateur système est un URI, par définition, alors que l'identificateur public est une simple chaîne. Traditionnellement, l'identificateur système d'une entité externe est un identificateur local ou propre à un système, alors que l'identificateur public est un nom permanent plus général. Les identificateurs publics s'intègrent mal à l'architecture Web existante, car ils ne sont pas des URI légaux. De nombreuses nouvelles spécifications (XSLT, XML Schema, etc.) ont pour exigence implicite ou explicite que tous les identificateurs externes soient des URI. Cet espace de noms a pour but de permettre le chiffrage des identificateurs publics en URN sous une forme fiable et comparable. Ce document décrit un processus permettant de présenter les identificateurs publics en tant qu'URN en introduisant un espace de noms d'identificateur public - « publicid ».

W3C/IETF URI Planning Interest Group. (2001). URIs, URLs and URNs : clarifications and recommendations 1.0. W3C Note 21, septembre 2001. Réference du 5 mars 2002.
URL : www.W3.org/TR/2001/NOTE-uri-clarification-20010921

Haut de la page

URNs et NBNs

Hakala, J. (2001). Using national bibliography numbers as uniform resource names. RFC 3188. Internet Engineering Task Force (IETF). Réference du 5 mars 2002
URL : www.ietf.org/internet-drafts/draft-hakala-nbn-01.txt

Hakala, Juha. URN implementation in national libraries. Réference du 5 mars 2002.
URL : www.lib.helsinki.fi/urn/urnimp.html

Hakala, Juha. URN Namespace ID Registration for the National Bibliography Number

(NBN). Réference du 5 mars 2002.
URL : http://www.lib.helsinki.fi/urn/NBN_registration2.htm Dead link - only reference found =
URL : http://lists.research.netsol.com/pipermail/urn-nid/1999-August/000014.html

Haut de la page

Autres espaces de noms

Cover, Robin. (2001). Namespaces in XML. The XML Cover Pages, 23 août 2001. Réference du 5 mars 2002.
URL : http://xml.coverpages.org/namespaces.html
Les espaces de noms XML « constituent une méthode simple pour qualifier les noms d'éléments et d'attributs utilisés dans les documents en langage de balisage extensible en les associant aux espaces de noms identifiés par les références URI. Un espace de noms XML est un groupe de noms, identifiés par une référence URI (RFC2396) et utilisés dans des documents XML comme types d'éléments et noms d'attribut. La spécification « Espaces de noms en XML » résout les éventuels conflits de noms au moyen de l'infrastructure d'adressage Web. La nouvelle spécification permet aux auteurs d'amalgamer deux ou plusieurs langages basés sur XML dans un même document sans entraîner de conflit ni d'ambiguïté, facilitant ainsi le développement modulaire et la réutilisation des langages et applications XML. Combinées aux espaces de noms, la modularité et la simplicité de la technologie XML ouvrent la voie à de nouvelles recherches, comme les travaux menés par le Groupe de travail sur le schéma XML (XML Schema Working Group) de W3C, et permettent l'échange de données à partir de l'architecture du Cadre de description des ressources (Resource Description Framework) élaboré par W3C. »

Haut de la page

Systèmes d'indentification des Éditeurs

American Institute of Physics. Publisher Item Identifier as a means of document identification. College Park, MD : American Institute of Physics. Réference du 5 mars 2002.
URL : http://www.aip.org/epub/piius.html

Elsevier Science. (1998). Publisher Item Identifier as a means of document identification. Elsevier Science Ltd. Oxford, UK. Réference du 5 mars 2002.
URL : http://www.elsevier.nl/homepage/about/pii
NB : Même notice que le précédent.

ISSN International Centre. ISSN (International Standard Serial Number). Paris, France. Réference du 5 mars 2002.
URL : www.issn.org:8080/pub/

International standard technical report number (ISRN) ISO 10444:1994. Réference du 5 mars 2002.
URL www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=18506
NB : Bon de commande et abstrait pour publication en vente seulement.

International standard music number (ISMN) ISO 10957:1993. Réference du 5 mars

2002.
URL : www.iso.org/iso/en/
CatalogueDetailPage.CatalogueDetail?CSNUMBER=18931
NB : Bon de commande et abstrait pour publication en vente seulement.

International standard book numbering (ISBN) ISO 2108:1992. Réference du 5 mars 2002.
URL : www.iso.org/iso/en/
CatalogueDetailPage.CatalogueDetail?CSNUMBER=6898
NB : Bon de commande et abstrait pour publication en vente seulement.

National Information Standards Organization (NISO). (1996). Serial Item and Contribution Identifier Standard. ANSI/NISO Z39.56-1996 (Version 2). Réference du 5 mars 2002.
URL : http://sunsite.berkeley.edu/SICI NB : édition autoriser disponible à:
URL : www.niso.org/standards/

National Information Standards Organization (NISO). (2002) Book Item and Component Identifier. Draft Standard for Trial Use. Bethesda, MD : National Information Standards Organization. Réference du 5 mars 2002.
URL : www.niso.org/pdfs/BICI-DS.pdf

Publisher Item Identifier Adopted. American Institute of Physics. Réference du 5 mars 2002.
URL : http://www.aip.org/epub/piipr.html
NB : Notice pour citation précédent.

Haut de la page

Permanence dans les citations de revues

LOCKSS project status - alpha and beta : permanent publishing : local control of content delivered via the web. Réference du 5 mars 2002.
URL : http://lockss.stanford.edu/projectstatus.htm

Reich, Vicky and David Rosenthal. (2001). LOCKSS: a permanent web publishing and access system. D-Lib Journal, juin 2001. Réference du 5 mars 2002.
URL : www.dlib.org/dlib/june01/reich/06reich.html
Article sur LOCKSS (Lots Of Copies Keep Stuff Safe), outil permettant aux bibliothèques de garantir à leur communauté un accès continu aux revues scientifiques publiées sur le Web.

Haut de la page

Ressource Australiennes sur la question de la permanence

Dack, Diana. (2001). Persistent identification systems, report on a consultancy conducted by Diana Dack for the National Library of Australia, mai 2001. Canberra, Australia : National Library Of Australia. Réference du 5 mars 2002.
URL : www.nla.gov.au/initiatives/persistence
PIcontents.html

Dack, Diana. (2001). Persistence is a virtue. A paper presented by Diana Dack, at the Information Online 2001 Conference, Sydney, Australia, 16-18 janvier 2001. Canberra, Australia : National Library Of Australia. Réference du 5 mars 2002.
URL : www.nla.gov.au/nla/staffpaper/2001/dack.html

National Library Of Australia. (2001). Persistent identifier scheme adopted by the National Library of Australia, septembre 2001. Canberra, Australia : National Library Of Australia. Réference du 5 mars 2002.
URL : www.nla.gov.au/initiatives/nlapi.html

National Library of Australia. (2000). Request For Quotation No. 00/99, provision of digital object management system for the National Library of Australia. juin 2000. Canberra, Australia : National Library Of Australia. Réference du 5 mars 2002.
URL : www.nla.gov.au/dsp/doms/domsrfq.html
Détails sur les exigences techniques et fonctionnelles et les critères de sélection de l'appel d'offres concernant le Système de gestion des objets numériques (Digital Object Management System) de la NLA, et réponses des fournisseurs.

National Library of Australia. (2000). Safeguarding Australia's web resources : guidelines for creators and publishers by National Library of Australia. Canberra, Australia : National Library Of Australia. Réference du 5 mars 2002.
URL : http://nla.gov.au/guidelines/2000/webresources.htm
NB: Édition juin 2002 disponible à
URL : www.nla.gov.au/guidelines/webresources.html

Haut de la page

Organisations

Internet Assigned Numbers Authority (IANA). Domain name services. Réference du 5 mars 2002.
URL : www.iana.org/domain-names.htm
Comprehensive source of information on Domain Name Services

Internet Corporation for Assigned Names and Numbers (ICANN). Réference du 5 mars 2002.
URL : www.icann.org

Internet Engineering Task Force (IETF). Réference du 5 mars 2002.
URL : www.ietf.org/overview.html

NISO - National Information Standards Organization. (2002). Bethesda, MD : National Information Standards Organization. Réference du 5 mars 2002.
URL : www.niso.org
Vaste source d'information sur les services de noms de domaines.

World Wide Web Consortium (W3C). Réference du 5 mars 2002.
URL : www.w3.org

Haut de la page

OpenURL

OpenURL demonstrator - Distributed Systems - UKOLN. (2001). Bath, UK : UKOLN. Réference du 5 mars 2002.
URL : www.ukoln.ac.uk/
distributed-systems/openurl

NISO - National Information Standards Organization. (2001). NISO Creates OpenURL Standard Committee NISO, mars 2001. National Information Standards Organization, Bethesda, MD. Réference du 5 mars 2002.
URL : www.niso.org/news/releases/PR-OpenURL.html

Open Name Services OCLC. Réference du 5 mars 2002.
URL : http://names.oclc.org/openurl

Powell, Andy et Ann Apps. (2001). Encoding OpenURLs in Dublin Core Metadata. Ariadne, Issue 27, mars 23, 2001. Réference du 5 mars 2002.
URL : www.ariadne.ac.uk/issue27/metadata

Van de Sompel, Herbert, Patrick Hochstenbach et Oren Beit-Arie. (2000). Open URL Syntax Description. Réference du 5 mars 2002.
URL : www.exlibrisgroup.com/sfx_openurl_syntax.htm

Haut de la page

Divers

Byrnes, Margaret. (2000). Assigning permanence ratings to NLMs electronic publications. National Library of Medicine, 2000. Mountain View, CA : RLG. Réference du 5 mars 2002.
URL : www.rlg.org/events/pres-2000/byrnes.html

Byrnes, Margaret. (2000). Defining NLMs commitment to the permanence of electronic information. ARL : a bimonthly report on research library issues and actions from ARL, CNI, and SPARC. 212 : octobre 2000. Réference du 5 mars 2002. Margaret Byrnes, 2000.
URL : www.arl.org/newsltr/212/nlm.html

Mealing, M. (2002). Dynamic Delegation Discovery System (DDDS) part one: the comprehensive DDDS standard. Internet Engineering Task Force (IETF). Réference du 5 mars 2002.
URL : http://search.ietf.org/internet-drafts/draft-ietf-urn-ddds-toc-03.txt
Ce texte précise les documents composant la norme Dynamic Delegation Discovery System (DDDS). Le DDDS est un algorithme abstrait permettant d'appliquer des règles de transformation de chaînes extraites par voie dynamique à une chaîne propre à une seule application. Ce document ainsi que RFC XXXX, RFC YYYY et RFC ZZZZ rendent obsolètes les documents RFC 2168 [8] et RFC 2915 [6] et mettent à jour RFC 2276 [5].

Haut de la page

Rapport - Table des matières

Création : 1996-05-22
Mise à jour : 2004-07-05

Haut de la page Avis importants